

1 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTAAACC TCATCAAATA
51 TGGCATCTCC CTTGCTTGCT GCAGCAGGGA TGGGAAGAAAT GTCACCTTTCT
101 TTTTAAGCTA GCAAGCTTTT TCTTTTCTT TTTCTTCTTC TATTTAAAAA
151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCCTTA TGACGGGGGA
201 GGAGACAATA TTCCCCTGAG GGAATTACAT AAAAGAGGAA CTCATTATAC
251 AATGACAAAT GGAGGCAGCA TTAACAGTTC TACACATTTA CTGGATCTTT
301 TGGATGAACC AATTCCAGGT GTTGGTACAT ATGATGATTT CCATACTATT
351 GATTGGGTGC GAGAAAAATG TAAAGACAGA GAAAGGCATA GACGGATCAA
401 CAGCAAAAAG AAAGAATCAG CATGGGAAAT GACAAAAAGT TTGTATGATG
451 CGTGGTCAGG ATGGCTAGTA GTAACACTAA CAGGATTGGC ATCAGGGGCA
501 CTGGCCGGAT TAATAGACAT TGCTGCCGAT TGGATGACTG ACCTAAAGGA
551 GGGCATTTGC CTTAGTGCGT TGTGGTACAA CCACGAACAG TGCTGTTGGG
601 GATCTAATGA AACACATTT GAAGAGAGGG ATAAATGTCC ACAGTGAAAA
651 ACATGGGCAG AATTATCAT AGGTCAAGCA GAGGTCCCTG GTTCTTATAT
701 CATGAACTAC ATAATGTACA TCTTCTGGGC CTTGAGTTTT GCCTTTCTTG
751 CAGTTTCCCT GGTAAGGTA TTTGCTCCAT ATGCCTGTGG CTCTGGAATT
801 CCAGAGATTG AACTATTTT AAGTGGATT ATCATCAGG GTTACTTGGG
851 AAAATGGACT TTAATGATTA AAACCATCAC ATTAGTCCTG GCTGTGGCAT
901 CAGGTTTGAG TTTAGGAAAA GAAGGTCCCC TGGTACATGT TGCTGTGTC
951 TCGGAAATA TCTTTTCTTA CCTTTTCCA AAGTATAGCA CAAACGAAGC
1001 TAAAAAAGG GAGGTGCTAT CAGCTGCCCTC AGCTGCAGGG GTTCTGTAG
1051 CTTTGGGTGC ACCAATTGGA GGAGTTCTTT TTAGCCTGGA AGAGGTTAGC
1101 TATTATTTT CTCTCAAAAC TTTATGGAGA TCATTTTTTG CTGCTTTAGT
1151 GGCTGCATTT GTTTTGAGGT CCATCAATCC ATTTGGTAAC AGCGCTCTGG
1201 TCCTTTTTTA TGTGGAGTAT CATAACCAT GGTACCTTTT TGAAGTGT
1251 CTTTTTATTC TTCTAGGGGT ATTTGGAGGG CTTTGGGGAG CTTTTTCAT
1301 TAGGGCAAAT ATTGCCTGGT GTCGTCGACG CAAGTCCACG AAATTTGGAA
1351 AGTATCCCGT TCTGGAAGTC ATTATGTGTT CAGCCATTAC TGCTGTGATA
1401 GCCTTCCCTA ATCCATACAC TAGGCTAAAC ACCAGTGAAC TGATCAAAGA
1451 GCTTTTTTACA GACTGTGGTC CCCTGGAATC CTCTTCTCTT TGTGACTACA
1501 GAAATGACAT GAATGCCAGT AAAATGTGCG ATGACATTCC TGATCGTCCA
1551 GCAGGCATTG GAGTATATTC AGCTATATGG CAGTTATGCC TGGCACTCAT
1601 ATTTAAATC ATAATGACAG TATTCACCTT TGGCATCAAG GTTCCATCAG
1651 GCTTGTTCAT CCCCAGCATG GCCATTGGAG CGATCGCAGG AAGGATTGTG
1701 GGGATTGCGG TGGAGCAGCT TGCCTACTAT CACCACGACT GGTATCTT
1751 TAAGGAGTGG TGTGAGGTCT GGGCTGATTG CATTACACCT GGCCTTTATG
1801 CCATGGTTGG TGCTGTGCA TGCTTAGGTG GTGTGACAAG AATGACTGTC
1851 TCCCTGGTGG TTATGTGTTT TGAGCTTACT GGAGGCTTGG AATATATTGT
1901 TCCCCTTATG GCTGCAGTCA TGACCAGTAA ATGGGTGGA GATGCCTTTG
1951 GCAGGAAGG CATTATGAA GCACACATCC GATTAAATGG ATACCTTTTC
2001 TTGGATGCAA AAGAAGAATT CACTCATAAC ACCCTGGCTG CTGACGTTAT
2051 GAGACCTCGA AGGAATGATC CTCCCTTAGC TGTCTTGACA CAGGACAATA
2101 TGACAGTGA TGATATAGAA AACATGATTA ATGAAACCAG CTACAATGGA
2151 TTTCCGTGCA TAATGTCAAA AGAATCTCAG AGATTAGTGG GATTTGCCCT
2201 CAGAAGAGAC CTGACAATTG CAATAGAAAG TGCCAGGAAA AAACAAGAAG
2251 GTATCGTTGG CAGTTCTCGG GTGTGTTTTG CACAGCACAC CCCATCTCTT
2301 CCAGCAGAAA GTCCTCGGCC ATTGAAGCTT CGAAGCATTC TTGACATGAG
2351 CCCTTTTACA GTGACAGACC ACACCCAAT GGAGATTGTG GTGGATATTT
2401 TCCGAAAGCT GGGACTGAGG CAGTGCCTTG TAACTACAA TGGGCGCCTC
2451 CTTGGCATTA TAACAAAAA AGATATCCTC CGGCATATGG CCCAGACGGC
2501 AAACCAAGAC CCCGCTTCAA TAATGTTCAA CTGAATCTCA CAGATGAGGA
2551 GAGAGAAGAA ACGGAAGAGG AAGTTATTT GTTGAATAGC ACAACTCTTT
2601 AACCTGAGGG AGTCATCTAC TTTTTTTTCC TCCTTTACAA AAAAAGAAAG
2651 GAAATATAAA AGCCGGGTTT TTGCAACATG GTTTGCAAAT AATGCTGGTG
2701 GAATGGAGGA GTTGTTTGGG GAGGGAAAGG AGAGAGAAGG AAAGGAGTGA
2751 GGTATTTCCC GTCTAACAGA AAGCAGCGTA TCAACTCCTA TTGTTCTGCA
2801 CTGGATGCAT TCAGCTGAGG ATGTGCCTGA TAGTGCAGGC TTGCGCCTCA
2851 ACAGAGATGA CAGCAGAGTC CTCGAGCACC TGGCCTGTTG CTCCAACATT
2901 GCAAAGACAC ATTATCAGTC CCTATTTCTA GAGGGATTAC TTTGAATTGA
2951 GCCATCTATA AACTGCAAG GTCTTGCCCT TTTTTTTAAT CAAAAGTGT
3001 CTGTTTAATT CATGAATTGT ATAGTTAAGC ATTACCTTTT TACATTCCAG
3051 AAGAGCCTTT ATTTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTACTG
3101 AGCTGTAACA AAGCCTCTTT AAATCGGTGT ATCCTTTTGA AGCAGTCCTT

FIGURE 1, page 1 of 3

3151 TCTCATATTG AGATGTACTG TGATTTTACT GAGGTTTCAT CACAAGAAGG
3201 GAGTGTTTCT TGTGCCATTA ACCATGTAAGT TTGTACCATC ACTAAATGCT
3251 TGGAACAGTA CACATGCACC ACAACAAAGG CTCATCAAAC AGGTAAAGTC
3301 TCGAAGGAAG CGAGAACGAA ATCTCTCATT GTGTGCCGTG TGGCTCAAAA
3351 CCGAAAACAA TGAAGCTTGG TTTTAAAGGA TAAAGTTTTC TTTTGTGTTT
3401 TCCTCTCAGA CTTTATGGAT AATGTGACCG GGTCTTATGC AAATTTTCTA
3451 TTTCTAAAAC TACTACTATG ATATACAAGT GCTGTTGAGC ATAATTAAAT
3501 AAAATGCTGC TGCTTTGACA GTAAAGAGAA AAAAAAAAAA AAAAAAAAAA
3551 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3601 AAAAAAAAAA AAAAAAAAAA AAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-158
Start Codon: 159
Stop Codon: 2532
3'UTR: 2535

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

			Score	E
CRA 18000005109762	/altid=gi 2599548	/def=gb AAB95161.1 (AF029...	1575	0.0
CRA 18000005109763	/altid=gi 2599550	/def=gb AAB95162.1 (AF029...	1573	0.0
CRA 18000005227216	/altid=gi 4762023	/def=gb AAD29440.1 AF14277...	1572	0.0
CRA 18000004989660	/altid=gi 4502869	/def=ref NP_001820.1 chlo...	1570	0.0
CRA 18000005231972	/altid=gi 8134363	/def=sp Q9R279 CLC3_CAVPO ...	1561	0.0
CRA 18000004989700	/altid=gi 6680948	/def=ref NP_031737.1 chlo...	1560	0.0
CRA 18000004978791	/altid=gi 1705905	/def=sp P51792 CLC3_RAT CH...	1560	0.0
CRA 1000685681515	/altid=gi 6634696	/def=emb CAA71072.2 (Y0994...	1559	0.0
CRA 18000004989661	/altid=gi 1705903	/def=sp P51790 CLC3_HUMAN ...	1558	0.0
CRA 18000005226296	/altid=gi 4753144	/def=gb AAB88634.2 (U8346...	1556	0.0

EST:

		Score	E
gi 10993825	/dataset=dbest /taxon=96...	1562	0.0
gi 10934924	/dataset=dbest /taxon=96...	1336	0.0
gi 10952244	/dataset=dbest /taxon=96...	1251	0.0
gi 12383593	/dataset=dbest /taxon=96...	1205	0.0
gi 6591096	/dataset=dbest /taxon=9606 ...	1170	0.0
gi 10251711	/dataset=dbest /taxon=96...	1104	0.0
gi 2321385	/dataset=dbest /taxon=9606 ...	1045	0.0
gi 5594360	/dataset=dbest /taxon=9606 ...	975	0.0
gi 5422132	/dataset=dbest /taxon=9606 ...	965	0.0
gi 10327969	/dataset=dbest /taxon=96...	963	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

gi 10993825	Neuronal precursor cells-teratocarcinoma
gi 10934924	Whole embryo-mainly head
gi 10952244	Neuronal precursor cells-teratocarcinoma
gi 12383593	Small intestine-duodenal adenocarcinoma
gi 6591096	Lung-small cell carcinoma
gi 10251711	Breast-normal
gi 2321385	Schwannoma tumor
gi 5594360	Brain-tumor
gi 5422132	Testis
gi 10327969	Lung-large cell carcinoma

```

1 MDASSDPYLP YDGGGDNIPL RELHKRGTHY TMTNGGSINS STHLLDLLDE
51 PIPGVGTYYDD FHTIDWVREK CKDRERHRI NSKKKESAW MTKSLYDAWS
101 GWLVVTLTGL ASGALAGLID IAADWMTDLK EGICLSALWY NHEQCCWGSN
151 ETTFEERDKC PQWKTWAEI IGQAEGPGSY IMNYIMYIFW ALSFAFLAVS
201 LVKVFPAYAC GSGIPEIKTI LSGFIIRGYL GKWTLMIKTI TLVLAVASGL
251 SLGKEGPLVH VACCCGNIFS YLFPKYSTNE AKKREVL SAAAGVSVAFG
301 APIGGVLFSL EEVSYYFPLK TLWRSFFAAL VAAFVLR SIN PFGNSRLVLF
351 YVEYHTPWYL FELFPFILLG VFGGLWGAF IRANIAWCRR RKSTKFGKYP
401 VLEVIIVAAI TAVIAFPNPY TRLNTSELIK ELFTDCGPLE SSSLCDYRND
451 MNASKIVDDI PDRPAGIGVY SAIWQLCLAL IFKIIMTVFT FGIKVP SGLF
501 IPSMAIGAIA GRIVGIAVEQ LAYYHHDWFI FKEWCEVGAD CITPGLYAMV
551 GAAACLG GVT RMTVSLVIV FELTGGLEYI VPLMAAVMTS KWWGDAFGRE
601 GIYEAHIRLN GYPFLDAKEE FTHTTLAADV MRPRRNDPPL AVLTQDNMTV
651 DDIENMINET SYNGFPVIMS KESQRLVGFA LRRDLTIAIE SARKKQEGIV
701 GSSRVCF AQH TPSPAPESPR PLKLSILDM SPFTVTDHTP MEIVVDIFRK
751 LGLRQCLVTH NGRLLGIITK KDILRHMAQT ANQDPASIMF N (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 5

1	90-93	NETT
2	364-367	NTSE
3	392-395	NASK
4	587-590	NMTV
5	598-601	NETS

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

1	24-27	KKES
2	330-333	RRKS
3	331-334	RKST

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 8

1	22-24	SKK
2	333-335	STK
3	529-531	TSK
4	613-615	SQR
5	631-633	SAR
6	642-644	SSR
7	658-660	SPR
8	709-711	TKK

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 13

1	27-30	SAWE
2	34-37	SLYD
3	92-95	TFFE
4	93-96	TFEE
5	105-108	TWAE
6	217-220	STNE
7	249-252	SLEE
8	383-386	SLCD
9	589-592	TVDD
10	666-669	SILD
11	674-677	TVTD
12	679-682	TPME
13	709-712	TKKD

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 18

1	49-54	GLASGA
2	53-58	GALAGL
3	72-77	GICLSA
4	88-93	GSNETT
5	189-194	GLSLGK
6	206-211	GNIFSY
7	234-239	GVSVAF
8	240-245	GAPIGG
9	245-250	GVLFSL
10	310-315	GVFGGL
11	313-318	GGLWGA
12	314-319	GLWGAF
13	408-413	GVYSAI
14	447-452	GAIAGR
15	491-496	GAAACL
16	541-546	GIYEAH
17	638-643	GIVGSS
18	692-697	GLRQCL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	99	119	1.810	Certain
2	182	202	2.131	Certain
3	233	253	1.398	Certain
4	256	276	1.019	Certain
5	290	310	1.770	Certain
6	321	341	0.797	Putative
7	361	381	2.093	Certain
8	400	420	1.539	Certain
9	473	493	1.739	Certain
10	496	516	1.218	Certain
11	540	560	1.568	Certain
12	570	590	0.975	Putative

BLAST Alignment to Top Hit:

>CRA|18000005109762 /altid=gi|2599548 /def=gb|AAB95161.1| (AF029346)
chloride channel protein 3 [Homo sapiens] /org=Homo
sapiens /taxon=9606 /dataset=nraa /length=818
Length = 818

Score = 1572 bits (4026), Expect = 0.0
Identities = 764/765 (99%), Positives = 764/765 (99%)

Query: 27 GTHYTMNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 86
GTHYTMNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54 GTHYTMNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173

Query: 147 WGSNETTFEERDKCPQWKTWAEIIIGQAEGPSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
WGSNETTFEERDKCPQWKTWAEIIIGQAEGPSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174 WGSNETTFEERDKCPQWKTWAEIIIGQAEGPSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF 326
NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF
Sbjct: 294 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF 353

Query: 327 FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 386
FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA
Sbjct: 354 FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 446
WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD
Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 473

Query: 447 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMIAI 506
YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMIAI
Sbjct: 474 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMIAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 626
VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL
Sbjct: 594 VVIVFELTGGLLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 653

Query: 627 AADVMPRLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFA LRDLT 686
AADVMPRLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFA LRDLT
Sbjct: 654 AADVMPRLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFA LRDLT 713

Query: 687 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD 746
IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD
Sbjct: 714 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDTHTPMEIVVD 773

Query: 747 IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774 IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 818 (SEQ ID NO:4)

>CRA|18000004989660 /altid=gi|4502869 /def=ref|NP_001820.1| chloride
channel 3; ClC-3 [Homo sapiens] /org=Homo sapiens
/taxon=9606 /dataset=nraa /length=820
Length = 820

Score = 1567 bits (4013), Expect = 0.0
Identities = 764/767 (99%), Positives = 764/767 (99%), Gaps = 2/767 (0%)

Query: 27 GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 86
GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54 GTHYMTNGGSINSSTHLLDLLDEPIPGVGTYYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114 SAWEMTKSLYDAWSGWLVTTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173

Query: 147 WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174 WGSNETTFEERDKCPQWKTWAEIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF 326
NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF
Sbjct: 294 NIFSYLEPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLSEVSYFFPLKTLWRSF 353

Query: 327 FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 386
FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA
Sbjct: 354 FAALVAAFVLRSLNPFNGSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGGAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 446
WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD
Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 473

Query: 447 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMIAI 506
YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMIAI
Sbjct: 474 YRNDMNASKIVDDIPDRPAGIGVYSIAIWQLCLALIFKIIIMTVFTFGIKVPSGLFIPSMIAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE--EFTHT 624
VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE EFTHT
Sbjct: 594 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHT 653

Query: 625 TLAADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYPNGFPVIMSKESQRLVGFALRRD 684
TLAADVMRP RNDPPLAVLTQDNMTVDDIENMINETSYPNGFPVIMSKESQRLVGFALRRD
Sbjct: 654 TLAADVMRPRRNDPPLAVLTQDNMTVDDIENMINETSYPNGFPVIMSKESQRLVGFALRRD 713

Query: 685 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDHTPMEIV 744
LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDHTPMEIV
Sbjct: 714 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSILDMSPTVTDHTPMEIV 773

Query: 745 VDI FRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
VDI FRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774 VDI FRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 820 (SEQ ID NO:5)

```
>CRA|1000685681515 /altid=gi|6634696 /def=emb|CAA71072.2| (Y09941)
    putative chloride channel ClC-3 [Xenopus laevis]
    /org=Xenopus laevis /taxon=8355 /dataset=nraa
    /length=791
    Length = 791
```

```
Score = 1559 bits (3993), Expect = 0.0
Identities = 745/791 (94%), Positives = 771/791 (97%)
```

```
Query: 1 MDASSDPYLPYDGGGDNIPLRELHKRGTHYTMNGGSINSSTHLLDLLDEPIPGVGTYYDD 60
        MD SSDPYLPYDGGGDNIPLR+LHKRGTHYT+TNGG+INS+THLLDLLDEPIPGVGTYYDD
Sbjct: 1 MDISSDPYLPYDGGGDNIPLRDLHKRGTHYTVTNGGAINSTHLLDLLDEPIPGVGTYYDD 60
(SEQ ID NO:6)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
CE00039	CE00039 chloride_channel	1671.9	0	1
CE00420	E00420 CLC	1288.1	0	2
PF00654	Voltage gated chloride channels	1172.4	0	1
PF00571	CBS domain	78.1	7e-20	2

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00654	1/1	71	622 ..	1	621 []	1172.4	0
PF00571	1/2	645	690 ..	11	54 .]	31.4	5.8e-07
CE00420	1/2	32	697 ..	1	729 [.]	1174.4	0
PF00571	2/2	726	778 ..	1	54 []	47.4	2.2e-11
CE00420	2/2	722	791 .]	867	942 ..	110.6	6.5e-32
CE00039	1/1	60	791 .]	1	804 []	1671.9	0


```

1 AATTCTATAC AAATATAATT ATATAGATAT ATATTACATA TACACACAAT
51 TGTTTATCTT TAAAAATAAT TCAAATATGG CTACAAAAC TTTACAATAT
101 GAAGCATTGT CAGTATTTAT TTTACCGGGA GGATTTCCTT CATCAGTGAG
151 TGCTGACTGT CATTTCATT CTTTATGATC AAGTTGTAGA TCAGGAAAAA
201 CAAGTTAAGA GAGTGCCTAC AAATACCGGG AAAACTTGTG GATAGATTTT
251 CATTTCCTAT GTAAAGACAT ATAAGAACAT GAATGGTATA AAAACAAAAT
301 CCTTTATAAA TGCCATACAA TTATATATTT AGAAAAATTA TATGGTGGTA
351 AAACATATAA AAGAACCACA CACTCCCAA TTTACATTGA GCTAATTTAG
401 TACAGTTAGC CTTTGTCAAA GCTTTCCTTG TTTAAAAAAA CTATTGGCTC
451 AGTGTGCAGG AAGGAGCATA GGAGAAAAAA TTGCCAAGAA TATTTGAAAA
501 ATACAGAAAA TAAAGAAAAA AATCACCTAC TATCCTATCA AAAATTTTAA
551 TAGCTAGAAT CAGGATAAGA TAGAATATTC CTGTGCAGT AATTCTAGTC
601 TATATTCCTT TCCTGGAACC CTGTCTCCCA AATTCAGGT GAGATTTTAT
651 AAGAAGCTCT GTTTATCTGA GATTTAAAAT ATAAAACTT GATTTAACCT
701 ATACAGTTT TTTAAAAGAC CCTAAATAAG TAAAAATTAG TACTCCACAA
751 ATTGAAGAGA ATTTCTCTCT TCTCTTACT GCCCTCTGAG TTTTCTCTTT
801 CCTTCTCTCA CCTCCAATTT TCATGTAAAC ACTTTCAGTT CGAGTGGACC
851 TTAGAGATTG TCTCATCAAA TACTTTAGGA AAACAAATTT TATAGAACCC
901 TTGAGTTCTG TGGAATTGCT TCTAATGAAC AACACCTTTT GTTGTGTGTT
951 TTGTTTAGTG AACTGTGTGA ACAGGCATTT CAGGAGGAGA ATCTCCAGT
1001 CTAGAGGAAT CCTCTCAGAG GTAGCTATAA AATATTGAAC TCTGATCTTC
1051 AATAAGCATT GTGCGGTTTT TGTTTTTGTT TTTAATGACA GTTTTAAACA
1101 AGAAAGTTGC TTTATTTCTG AACTTCATAA AAATTTCTAT TAAAGAGACA
1151 ATTTCTGAAT TTTTAAACAA TTTCTAGAAC AGTTGAGTAC CTCACTTTGA
1201 GACACATTTT TGCTAAAAGT TAAAAACACA AAACCCTTAT GAGATAAAAT
1251 AGGAAGCTAG TAGAGATAGG AAAGTCCTCT GCTTAGTAAA CCTCTTTTTT
1301 GCGTAGTTTA GACACATACA ATAGTAAAGT TACTTAGTAC GTTGATAGTT
1351 TTCTTTCTCC TCAAAAGCTA CAATGTCTTA CTAGCTAGTT CCTTCAAGAA
1401 AGGAAACAAG AAGCCGCTGG AGGAGATTGG TGAGTGGGAT AAAACACTAT
1451 TCAACTCTTC AGTTATTCGG TTTTAAATC CTCAATGAAA GGCTGCTGTA
1501 TTATAGAGTA TTTTTTTTTT ATTTTAAATA GACTTAGAAC CAAGTTTCTT
1551 GAGAAACCTT TGGCATATTG TAGTTTTTTT ATGGCTATGA CTCACATGAC
1601 ATTACTGTAT AAAACTAGTA CATTCTCTCG TAAAACCACA CAACTTACT
1651 AGAGTGCTGC TCTCATTTTT CTACATTAGA AATGAAAAAG GGCATTGTCT
1701 GCATTCAAAA TTTCTTTTTT ACATCTCTGT ATTACTTTTT CCCCTTTATA
1751 TTTATCTTAA AACCAAAAGA AATAATGTTT CTATTGTTTT ACTGTAGTTA
1801 CCACTGATGC TACCGAAGCT GTATTGTGAG TGTTTCAAAA TTCTCAAACC
1851 AGTTTTGTGT GTTGTACTTG TTTGTACTGC GAGCTTAGTC ATTGTCATAC GTAGCAGGAC
1901 CTGATTAGA AGGCTGTGCC GCCTCTAAGC CTTGCTAGAT TGTAGCCACT
1951 AGCAACCAGG CTGCAATAAT TTCCCTTTGA TGACATCATC CACTGTGGAA
2001 GAACCCAGTT GCTTCAGCGA GTCGAAC TAC AGTTTAAACC TCATCAAATA
2051 TGGCATCTCC CTTGCTTGCT GCAGCAGGGA TGGAAGAAAT GTCACCTTCT
2101 TTTTAAAGTA GCAAGCTTTT TCTTTTTCTT TTTCTTCTTC TATTTAAAAA
2151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCTTA TGACGGGGGA
2201 GGAGACAATA TTCCCTGAG GGAATTACAT AAAAGAGGTA ATACTATCCC
2251 CTTGCTGTGA ATTCTCTGTT GGTATGTTTT GCATGCGGCT GGGCGGTCCT
2301 CTAGCTTAAA CTGGTTCTCG TTTGTCTTTT AAATACTGCA GTACGTTGTT
2351 TAGTTGCCCT GGGTTGTTAG TAAGGGGAAA ATGCAACCTT CTGAATGGTT
2401 GTGTAGCCAT CCCTGATTGT TTTCTCTGTG CAGATTAGTA CTGCTTCAGA
2451 TCACGTCGGG CTCCGACTCC ATCTTCTGCA TGAAAACTT CTTTCTAACT
2501 CTGAAAAATGA ATTAATCTGC TTTTACAGCC AACTAAAGTC GTGTGGTTG
2551 GCATCTAAAA AGTAATGTTT TTCTTCCTTC AGAAAACTTA CATTTCCTTT
2601 AATTTACACA GAGAAATCAG GTGCCTATGT ACCATTATAT TTTAGCTGCT
2651 GCCAATTACC ATGTAGATTT TACACCACAA AGTAAATTTA TAGCAAAAGC
2701 TTTACCTACA TTTTAGAACA TTTTAAATG ATAGTAAAGA TGAATAATTT
2751 CTATATTAAT ACTTTTTATT TAATATGTAT TTCGGCTGAG TAACATACTA
2801 CATTGTCTCT CACAGGTATC TTGTGAAATT TGATATGATA AAACACATTT
2851 GACTAAATGT CAGAAAAAAT AATATTGGTT TGTGAAAAGC AGAAGAGCAC
2901 CCAGCATGCC TGTAATCTTT TTGGCAGGCA CTTCTCAGT CTCCTTAAAA
2951 TTAATTGCAT GTTAATTACT ACCCTTTTTT TCATTTTTGT TTAATTGCTT
3001 ATTCGAAAAA CAGACTGGTC GACATTTGTT GTCCTAGAAA AAAATTGAAC
3051 TTCAAGAAAA ATCTCTTAGC TTATGTGACT TCATTTTTGA GCCACATTAG
3101 TTTGAATTAC TGCATGATAT TATAAACTCA CCTTATGATT TAACCCAAAC

```

FIGURE 3, page 1 of 27

3151 TTTTATTTGT AAGTATATAA GGAAGTAATA ATGTTTTTCT AATATAATTA
3201 GCCTGCTTTA TTTAAAATAT ACTTTGTGTT CTGATAACAC TTTTTTTTTA
3251 GTATTAAGTT CCACTATAAT TTAAACATTA TAATGTATTC AACAAATGTC
3301 TGTTGGTTGC ATTGTGTCTG CTACACACTA TTTTAGGGTC TGAACAGTTG
3351 TAGCATTATT TATCTTGACG TATTCTGTAG TTAGTAAAAA CTTGCTTTTT
3401 ACATTTTGAG AAAAGCTGTG TAAGGATCAT GTTACATACA TTGTGCTTTC
3451 TCTTACAGAG TTACCTTCTT AATAAAATTT TGATATATGT GTATATGTAT
3501 ATGTTAGAAC ATTTGGAAGA AATATCTAAA AGCATAAAGA AGAAAATAAT
3551 TTCTTGTAAT CACACCACCC AGAGCTTTTT AAATTTTTTT TCTTAATGTT
3601 ACGATCATAA ATTCTTCTAT TTCCTATGTT CTGATTATCA GTTTTCTGGT
3651 AAGGATTCTT TTAACAGGA AGCAAGGTGA ATGAATAGTG ACTGTTCAAA
3701 TGTCACATTA TTTGCTAATC AGTAATTAAA CTGTAAAAA AGACAGACTG
3751 TATTTTCCCTC ATGCTATTAC AACATTTGGT TGTTAATGAT GATAGATCAG
3801 AATACCTGGG CTTCAGAAAT TTAAATTCCT TTTGTGAAGC TTAACAGTCT
3851 TTGACAGAAC TTACTTATGG ACTGTCTTAG TGTAATAATAT GCAAATAATA
3901 AGAAATAAGT CAAAACCTAT GTGAGAGTAG GCATGGTTAC TGATATTACC
3951 TAAACGTAAG CTTTTTATTT CTATTATACT TTCATAAATA ATCCTTTAAG
4001 AATCTTGCTT AGGATCTAAA TCAGTCCAC TCTTGGCAGC TCAAATAGGT
4051 TCTTTATCCC TTGATGAGAC TTATTCTATT AATATAAGTC ATTGTTATTT
4101 GAAAGTAACA TTGTGTATGT GTAGTAGAGA TAAGTCAGTT ATTAGGCTTT
4151 CGTGACTGTA CTGTATTACC TCAAACATAC TGTAGTATCC TAGTGTCTAT
4201 GCGTAAGATG TTATTTTTTG TCCATAATTT ATGACCTGTT GTAGCCATGG
4251 GTCAACACAA TGGAATTGAT GGAGACAGGC AGCTAACAAA TCGAAAAAAC
4301 TGAATCAGT TCCCTGTGAG GAAGAACAAA ACTATAATGA TTAATTTGA
4351 TCTTCAGCCT GATAGTGAAG AGGCAGATAA AGTATAAAAT TGTGAAGGAT
4401 ATCAATAAAG TAAACATGGA TCTGTTTAGT AAATCCCTGA GTGCTATAGC
4451 CAAGGATTAC CTTTGTGAG TAAATTGAAT TTAATACTAC TTTCAAGGC
4501 GAGATGGTAA ATGGTGAAGC TTCCTATTTA AGTAAATAAT GTCAAGTCTG
4551 GAAGTATAAG TAGATCAAA TTAGAATTAG TTTGATATAC TATTGATAGA
4601 TTAGAAATTA AGATGACATT TCAGAAATAG CCATCTTAG GGGTAGATTT
4651 CCTATATAGA AACAATCAAG CTCTCTCAA ATGTCTCTTC CTTTTTTATC
4701 AGGAAAAAAG ACTTGGCTTA TCTGGACTGT TAGTTTTACA CTTTTTCTTC
4751 TTAATTTGTT CAAGATGTTT AAGTAGTTTT AGAGGTCAA TTTCTTTCTT
4801 CTACCAACCC TTTATAATGG ATTTGATTCT TTTGGGCTG AGCCTCCATT
4851 TACTCCATGA GGGGCTTTA ACAATTATTT AAATNNNNNN NNNNNNNNNN
4901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5101 NNNNNNNNNN NNNNNAAAT AGTAATATTA ATAATAGTTA ATATTTATTA
5151 GAATTTCTCG TTAGCTGGAT ACTGTCCCTA AGTGGGTTT TTTGTTGTTG
5201 TTGTTGTTGT TGTTGTTTTC TTAAGAGAGA GGTATCACTT TTTACCCAG
5251 GCTGGAGTGC AGTGGAGTGA TTATAGCAA TGCAGCCTG AACTACTGGG
5301 CTTAGATCCT CCGTCTACC CTCTTGGA CTTGGGACTG CAGGCTTGCA
5351 ACACCTTGCC TGGCTAATTT AAAAAACAAA ATTTTTTTTT TTTTAGGGA
5401 GAGTCTCACT ATGTTGTCCA GGCTGGTCTC CAACTCCTGG GCTCAAGCAA
5451 CCCTCCTGCC TTGGCCTCCC AAGTAGCTGA GATTACAGGT GCGAGCCACT
5501 GTGCCTGGCT TGTTCTAAGT GCTTTATGTG TATGAAATTA TTTAAATCCT
5551 CATCACAAGT TTATGAAGTA GGTACTGTTA TAATCCCCAT TTTCTAGTTG
5601 ACAAGACTGA GGTAAGGAAT TGTTAAGGAA AAGTCAGAA TCCATCCAGA
5651 TATTTGGCTC ATACTTTAAT CATGAGGCTA AACTGCTTCT CTCTACCGT
5701 ATCTTCATAG TAACCTGTGT TTTAAGTCTG GTAGAAGCAT AAGAAGTTTA
5751 AACACAGACA GAATCCTGTG GAAGTTAGTA AATTTCTAGT GAACGATAGA
5801 AATGATAGAA ATCTCTCTT CCCCCAAAGT CCAAGAACA GATTAGTCTG
5851 CTTTTGACAA GTGTTATCAA AGTAGACTGT TCTCACATAC ACGGGGGACT
5901 CAATAGGGCA TTCCTGGTGG ATATAATAAA ATGAGTAAAT GCGATAACAG
5951 GAGGAAATGC CTAGTGTGTT GCTCTTGGAT TAGTTTTGAT ACAACAAAGG
6001 CAGCTTTGTT GTGAGTCAGT AGAGAGGGTA GTGTAGAAAG GTGGAAGTTG
6051 GAAGAGTGGC AGATCCTAGA GGACTAATGA TGGGCTTAAA CCACAAAAAG
6101 TGTCGCTTTG CCATTGAAAT AAAAGTTTGG GGTCTTATTT TTTCAATTTT
6151 CTCCCTGAAA TTATTTCTTG ACATTCATTA GCTCAGCAGT GTATCTAAAT
6201 AAAGCTTTTT TGGGTTTCTA TTATAATAGA GGTTTGTTC TTTTCTTCC
6251 CTTTGAAGAG TATCATTTTT TGCACATTAT TTGAAAATCC AGGTGTTATA

FIGURE 3, page 2 of 27

6301 TGATATTCTT ATTGCCAGAG GGACATTCTG CAGGCTCTTT GTAAAATGAT
6351 TTTAGGATTC AGATACTTAT TATATTTTAA TTGGCCCTAA TATTTTATCC
6401 AACTAGAAAA TTAAACCTCT TCTTAAAAAT TAATCCATCT AAGTGTCTGT
6451 AAATTAAGG AACAACTAAA GATTCTTTAT TTGGTGTGAG AAACCTCTTG
6501 TTTCTACAAC AGTAGTATAA AACAAAGCCT GTTTTTAAAT GTACTTTTCC
6551 CACAGTATCT GAATTTCAAA TCTTCAATAA AATCTGGTTC ATATTACTAC
6601 CTCTAGCTTG ATTTTCTAAA AATAGCTGAC ACTTTAGTAT GGTTAATTTT
6651 ATGCCATCTC ATGGCTTGTC AGAAATGCTT TGTATCAAGA TTTCGAGTG
6701 TGAACAGATT TCCTGCCGCA TTGATTAAGT TTGTAAATTT GGCTATTTTC
6751 CCAGCATCGA GGTTTCTGCT TTGCGTTTAT GCAGGAGACT GGTAGTTTAA
6801 ATTGAACCTT AAGGTTTGT TTCTTGTTT TAAGTTAACA TATGTTAAT
6851 TTCTAGTTTC TTTGTAGCCC TTTGCAACTT TAATTAGGTC ATAAAATGGA
6901 TTTACTCTAG TTTCTCTAAC AAATTTTATA AATTTATGAA ATATGAAATT
6951 TAGCAAATTT TATAAACCTT TTTATTCATG TATTGTACAG CTCATCATAT
7001 TTGCAGACAT AATAATGAA TGTGGAACCT GTTTCCAATT ACACAGATGT
7051 CTTAATATCC ACCTTATCAT CTCTAACTAA AGGATGTGGC TTTTATTTT
7101 TGAGGTGGCA ACAGAACAGA AAAGAAAACA GTGAATTGAG TAATGGGCTT
7151 AGTATTGCTG CTGCCCTGTT GTGTATCTTT GGTAAACTTC TTTGAGATTT
7201 GGCATTAAC TGCAAGCTT TGCAGTTTAG ACAGTTAAAT ATGACTGAAT
7251 GGCTGAACAA ATTTTAATAG CGTATGCTTC TTTTTTGCTA TTTATTTACC
7301 CAGTAGACAT TTAATTGACC ACCTGCTAAA TGTGAGGCAC TATTCTTGCC
7351 ATTACCTTTT TAATCTTTGA TTTGGAGTCT GCTAACATTC TGGAACTTCC
7401 ACTATCAACT TAGAACGTTT ACTTTCCCAT CCCTTACCAG GATGGCCATT
7451 TCTTATCAGT AGGGTCACAG AGAGAGAAAA AAAAAACCAG CTGGGGCTAG
7501 ACTTCCCTGCT CTTAACATAC AGAAGCAAAT AGGTGTGAA GGAATACATA
7551 GTATTTTGGG TTTCTGCCTC TTCCCTCCAT AATTTTTTTA AAAAGGTTCA
7601 TATGTTTTAT GTGTGCTTA TGTAACAGTA ATCTGCATTA TGAACATAA
7651 TGACGAGGAT CACCATTTCA CATCTTTGGA GATTGATCAC AGAGGTAATA
7701 AGTAACTCTT TTTAAATAAC TATATGCATC ATTTTTCATG TAAACTATT
7751 ATTTGGGATA ACCCTTTTGA GAAAAGGCTT AGGCTCCTGC CAGTGTCACT
7801 TGATATTTAA CTAATAAGCT CAGTTTAAGG CGCAGCAATT AAGGTTGTGT
7851 TGTTTTTTTT TTTTAAAGTT CAGTTCAGCA AATATATGTG GAAAGCTTGT
7901 GGGTAAAATT ATATTTGTAT TTTTGGGAAA GCAGACAATT TTATTAATGC
7951 CTATATTTTT CTAGTTCAGT GTTTGTCAA CTTCAAGTTT TAACATGTTG
8001 ATCATGAAAC CAGTTGACTT GTGACCAGTA TTTTAAAAGG AAAGATTAAA
8051 AAAACAAAAT AAAATATCAG TATATACCAA GTAGTAAGAG TAAGCATTGT
8101 TTAATAAATC TTGGTTTTAT TTAAGTACAT ATCTATATAC TATGTCAGTG
8151 AGAAACATTT CTCCACTTCA TGTTTGAAAA ACATTTCAAA AGCTAAGAAA
8201 AAGTTTGAAA ACCTGTTTGT AAGTACACCT GGGGTAAAGG TACACCCTGT
8251 GGCATAAGAT GTCGGGAACA ACTGAGGGTA AGAATGGGGA TGCATTACTA
8301 TCGTAAACTT CTGCTAAAGC ATAAGGATGT GAGTGCTGGG AGCAAAGCAG
8351 TGCTCACCAC TTCTGCAATT TTCTATTGCA GCATTTTAAA TAATATGGGA
8401 AAAAGTGGAC TGCAACCAAA GGCAAGAGG GATGGTGATG GTGAAGGGTA
8451 AGATTGTATT TATTGTCCAA AGGCTAAGTG CATATACATA TGTGTTTGGG
8501 AGAAGGCATC ACCTAATAGT TCTTAACCTA CTCTGAGAGA AGGTGTGCTA
8551 CATTCTCTAA AGTATACATG TAAACCAACA ATGAAATTAT TTTAGTGAAT
8601 TGAGAATCAA AGTGCTAGAG TTTGAATCCC TGTCTACTA CTTGCTAGCG
8651 GTGTGACCTT GGGCTGTTT AACTCTTGAC ACCTGTTTTT CCAAATTTAT
8701 AAAGTGGAGA TAATAATATC TGTCACATTG TGTGTTGTG AGGATTATAT
8751 GAACTAATAT ATGTAATGTC CTGAGAACAA TGTCTGGTAC ACATTAAGTT
8801 AATTAATAAT AGCTGTTCTT ACTGTTATTA TTAGACATGA GCTAGATAAC
8851 AGTGGCTCTC ACATGTGAAA GATTATTTTA ATTCTGATGT AGTTCAGTTT
8901 ATCTATTTTT TTTATTTTTG TCCCTTTTGC ATTGATGTCA TATCTAAAAA
8951 ACCTGCCTAA CTCAGGATCA CAAAAATTTA CTCCTGTATT TTATAATTTT
9001 AGCTCTTTAG ATCTAGGATC CATTTTTAGC TAATTTTAT ATATGGTGTG
9051 AGGTAGGGGT ACGGTTTCAT TCTTTTGAC GTGAATAGCC AGTTGTCCCA
9101 GCATCATTTA TTCAAAAGAC TATTCTTTCC TCACTAGAAA AAATATTTCT
9151 TTAAAGAATA ATGAATCCTT TTTTTTTTCT TTTTAACCGC TGTACTCAG
9201 TTGGAAAAAG AATAATGAAT AATTTTAAGT AATTTTCTTA CAGGTAAATT
9251 TAAGTCTTTA TGTTTAGATT ACACATATTA GGAAATAATG GATTGTGATT
9301 CCATAGGTAT GCTTGATCTT TATAAAGTTC CCTGTCTCTG GAAAACTAA
9351 AATAAGGCAA AACAATCTTC TTAGTAGAGT TATTTTACA AGAAAGTTGC
9401 AAGCCAGTTT TAGTTCATCG ATTGGATAAT TTTTCTGCT TGCTGGAGGT

FIGURE 3, page 3 of 27

9451 ATTTTCAGTAT TGGTAATACC TGAACATGA GGATGCATGA ATGATGCATT
9501 TTAGGAATTT GTTCTGTGT CCATACCAGG CATAATGAAT TAAGTTATCT
9551 GTTAAAAATA CAGGATTTTT GCTCAATATA CAGTTGTAGA AGAACTCATT
9601 GTCCAAATTT TTAAGACTTT TTTTCTTTT TTTTGTGAG ATGGATCTCG
9651 CTCTGTGCGC CAGGTTGGAG TGCAGTGGCA CAACCTCCAC TCACTGCAAC
9701 CTCCACCTCC AGGGTTCAAG TGATTCTGCT GCCTCAGCTT CCCGAGTAGC
9751 TGGGGACTAC AGGCATATGC CACTATGCCC GCCTGATTTT TTTTAGTAGA
9801 GATGGGGTTT CACCATATTG GCCAGGCTGC TCTTGAATC CTGACCTCGT
9851 GATCCACCCG CCTCAGCCTC CCAAAGTTCT GAGATTACAG GTGTGAGCCA
9901 CCGCGCCCGG CCAGACATTT TTTTTTTTTT TTTTTTTTTT GCTGTCTTTG
9951 TCATATTGTT AGTCTTTTGG TTAAGCGATA TTATACTTA GTCATATGAG
10001 TAATATAATG CAACATGCTG AATTGTGTGT GTGAGAGGGG GTGTGTTTTT
10051 GTTTGTTATT TGTTTTTTAA ATAGAGATGA GATCTCACTG TGTTTCCAG
10101 GCTCCCTTGA ACTCCTGGGC TCAGATGATA TAGCCTCCTG CCACAGCGTC
10151 CTGATTAGCT GGGACTACAG GTGTGCACCA CTACACGTGG CTTTCTGAT
10201 GAAATTTTAA ATACCCAAAT ATTTGAGCAG AAATAATAGC TTGTGTTTAT
10251 TGTTTTTCTA CTATCTGTCA AGTATAGTAT TAAATGTTTT ACATAATTTG
10301 TCTCCAGTCC ACATACAATA CTCTAGTAGA AGTGGGTAAC AAAACCAAGG
10351 TACTCAAAGA GGTTAATAAG TAACTTGCGC TGGATCACAG AACTAACGGG
10401 AGGCAGGGCT GGAATTTGAC TCTAGGTCTT TCTGACCTCA AAGTGCAGTA
10451 AAGTCATGGA ATTTCTCTAC TAGGCCACCT GGAAGAAAAG TGATCTTTTT
10501 TCCAGTCTTT TTTGTTACTG TTTTTCAGCC AGGAGATAGT AGAGTTAGGT
10551 AGTAGAATAG TAGTCACTGG CATCCGGTAG TCAGCCCTCC AAAAAAGTTT
10601 TTGATTTTTT TTTTTTTTTT TGTCTTAAAC TTGGAAGCTA CTAACTTTCA
10651 GGTCACTACT TCTTATCATC CAAGAGCTGG ATATTTAGGT AGCAGAAACT
10701 ATGGAATTAT CCTAAGTCCT CTTGAAGCTT CAGCTGTAA AATTAATTGG
10751 TTCTGATTAA CACTGTGCTC AAGATTTACA TTTCTAGGAG CCACAGTTTG
10801 ATTGGTCTAA CTTGGATCTA TGTGTTTTCT TTAGCTGGGG AGGAGAAGGT
10851 ATCTTGATTG ATACCTTCAC CAGGACTGCA TGCAGTGAGG GACAGAAGTT
10901 TCCTTAAAAA AATTGGGTTT TGTATAGGA AGAAGGGGAA GGAGATACCA
10951 AGTGGGCAAA ACAATCAGGT TCTATTACAT AAATAATAAA CCTAATGTGA
11001 CGATAATAAA TGGATAATAT GATTATTTTA AGTTTGGAAA TATACCTGGT
11051 TATTAGTATT GGATATCTGG TAGTGGGGTT GGAGAAAAAG TCGAGAATAA
11101 GAAAAGACTT AAAATCGTAA AAATTAACCT GAAAAGAGGA TGGCTGAGCA
11151 GATACATATA TGTTAGATAA TGTTTATAAT GGCAAACCAA CCTGAAGATT
11201 TGTTTAAATT GTAGTATGTA GCCAGGTGTG GTGGTGCTTG CCTGCAGTCC
11251 CAACTACTTG GGAGGCTGAG GCAGGATGAT TGCTTGAGCC TAGGTTTGAG
11301 GCTACAGTGA GCTATGTTTC CACCCTGCT TCCAGCCTA GGTGGCAGAG
11351 CAAGACCCCA TCTCTAAAAA AATAAAGTAA AATGAATAAA TTATAATATG
11401 TTATGACAAA TATAGTTATC TGAAGTCACA GAAAATGTGC ATGTGCATTT
11451 AATGATGTGA AATAATTTTT AGGAAGTATG AATAAAAAAA TCAACTTTTA
11501 AGTGTGGCTA GTATGATCTT ACCTGTATCT CACTTATAGA AAATATAAAA
11551 GGCTGAAGCC AGTCACCAGT TTAATAGTTC TAACCTCTTG TTTACTTGAT
11601 TCCCTTTTTT CTCCTCCCCA GCAATCCTCA TATAGTTAGG TAAAGTTGGT
11651 TCTTCTCAG CTTGTTGCA GAAACCCCTA AGCCTTTTTA CTTAAAGCTT
11701 TTTGAAACCC AGAAACCCAT CTTTGAATT CAAAAGTTTT GACTGTTATT
11751 AGTCTTTTTG TATGTTTGTG GGCCGCATAA ATGTCTCCTT TTTATGAACA
11801 GAGAAGTGTC TGTTAATATA CTTTGCCAC TTTTGATGG GGTGTTTGT
11851 TTTTTTCTTG TACATTTGTT TAAGTTCCTT GTAGATTCTG GATATTAGAC
11901 CTATGTCAGA TGGATAGATT GCAAAGTTT TCTCCATTC TGAGGTTGC
11951 TTGTTTATTC TGATGATAGT TTCTTTTACT GTGCAGAAGC TCTTAGTTT
12001 AATTAGATCC TATTTGTCTG TTTTGGCTTT TGTGCCATT GCTTTTGGTG
12051 TTTTCAATCA GAAGTCTTTG CCAGTGCCTA TGTCTGAAT GGTATTGCCT
12101 AGGTTTTTCT GGTTTTGGGT TTTACATTTA AGCCTCAAAT CGATCTTGAG
12151 TTAATTTTTG TATAAGGTGT AAGGAAGGGG TCCAGTTCCA GTTTTCTGCA
12201 TATGGATAGC CAGTTTTCCT AGCACCATT ATTAATATTA AATAGGGAAT
12251 CCTTCCCCCA TTACTTGTTT TTGTCAAGT TGCTGAAGAT CAGATGATTG
12301 TAGATGTGTG GTGTATTTT TGAGGTCTTT GTTCTGTTCC GTTGGTCTGT
12351 ATATGTGTTT TGGTACCAGT ACTATGCTGT TTTGGTTACT GAGCCTTGTA
12401 GTATAGTTTG AAGTCAGGTA GTATGATGCC TCCAGCTTTG TTATTTTGC
12451 TTAGGATTGT CTTGGCCATA CGGGCTCTTT TTTGGTTCCA TATGAAATTT
12501 AAAGTAGGTT TTTCTAATTT TGTGAGGAAA GTCAATGGTA GCTTGATGGG
12551 AATAGCGTTG AATCTATAAA TTACTTCGGG CAGTATGGCC ATTTTCATGA

FIGURE 3, page 4 of 27

```

12601 TATTGATTCT TCCTATCCAT GAGCATGGAA TGTTTTTCCA TTTGTTTGTG
12651 TCGTTTCTTA TTTCTTGGG CAGTGGTTTG TAGTTCTCCT TGAACAGGTC
12701 CTTACGCTCT CTTTAAAGTT GTACTCATCA TCACTGATCA TTAGAGAAAT
12751 GAAAATCAAA ACCACAATGA GATGTCATCT CATGCCAGTC AAATGGTGAT
12801 TATTATAAAA AGTCAAAAAA GAATAGATGT GGGTAAGGCT GTGGAGAAAT
12851 AGGAATGCTT TTACACTGTT GGTGGGAGTG TAAATTAGTT CAACCATTGT
12901 GGAAGACAGT ATGGCGATTC CTCAAGGATC TAGAACCAGA AATACCATTT
12951 GACCCAGCAG TCCCATTAAT GGGTGTATAC CCAAAGGATT ATAAATCATT
13001 CTGCTATAAA GACACATGCA CACGTATGTT TATTATAGCA CTATTTACAA
13051 TAGCAAAGAC TTGAAACCAA CCAAAAAAGC CATCAATGGT AGACTGGATA
13101 AAGAAAATGT GGCACATATA TACCATGGAA TACTATNNNN NNNNNNNNNN
13151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14501 NNNNTAAAG ATACATCCTT TATTCATGCG TAAGATGAAA TCGAGAGGTG
14551 AAATTGGATA TACTGTTGCT TTTAAAAAAT TTAAACATAT ATGTAATTTT
14601 TTGTACTTAT CTCATTTTAG CCTATATAAG TTATATATAT TTTGTTTGTT
14651 TGTTTGTTTG TTTTGTTTGA GATGGAGTCT TGCTCTGTCA CCCAGGCTAG
14701 AGTGCAGTGG TGCAATCTCG GCTCACTGCA ACCTTCGCCT CCTGCATTCA
14751 AGCGATTCTC CTGCCTCAGC CTCCTGAATA GCTGGGATTA CAGGCACCTG
14801 CCACCGCGCC CAGCTAATTT TTTTATTATT TAGTAGAGAC AGGGTTTCAC
14851 CATCTTGGCC AGGCTGGTCT TGAACCTCTG ACCTTGTGAT CCATGTGCCT
14901 TAGCCTCCCA AAGTGCTGGG ATTACAAGCG GGAGCCACCG CGCCCGGCTG
14951 TAAGTTATAT CTTACACAAA TCTAGGTTTC ATTCAGAGAA TTATATGCAA
15001 AGAAACAGTG CAATAGGATT ATTTTAAAGC TATTGTTATT GTTAGAAAAC
15051 ATAATACCTT TAAAATTCCT TTTCACATTA GAAATATAGT GGCTTCTCCC
15101 CAGTTTAGGA TAGAAATTTT CCTTTTCTTC TCCTTCTTTA TACTATTCAG
15151 ATTTGCATGT TTGACAGAAC AAATTATAAG AGAAAATATT TGAAATGTCA
15201 CATACTAAAG TAAATGTTTG AATGTTTGAA AATTTTCTGG TTTTCAGAGA
15251 TTTTGAATTG CTGAATCGTT GTGTAAATTA AGATGTTGAG TAGTTTCCAC
15301 AGAGTAATTA TTTGAAAGTC ACTGAAAGCA AGACACATGC CTAATGTAAA
15351 TGTTTATTGC ACTACTGTAC CTTTTTCTAC CTCATAAAAA TGAGAATAGC
15401 AGTCTGTACT TTTCCACTTC GTCATTCGTA AGTCTTTGCA GAAATTCATA
15451 TTTTGTTTGC TTATTATCTT CACGCTGTAA ATAGCTTGAA AATTCTTTAA
15501 GTGGGGCTAG CGATGTATTA TGGATACATG TTAAGTGGTA TAGAAATTTC
15551 ACTTTTTTTT TTTTGCATAA AGAGTAACAA GACCAGTAGT CCATATTCTT
15601 TCAGCTCTAC CCAGAGAAGG GCAATGTAGG AGGGAAATG AAGTTTGCAA
15651 AATATTTCAT AGTAGGCTTT TTCTTAAAGT AACTTCAGAC TTACAGAAGT
15701 TAAAAATAG TACAAAGAAT CCCCATATAC CTGTCACCCC AATTCCTGAA

```

FIGURE 3, page 5 of 27

15751 ATATTAATAT TTTACCACAT TTGTTTCATTA TGTCTGTATT CTCCAAGTAC
15801 GATATATGCC ATTATATGTA ATATGTAGCA TTTTATATAG ACATAGGGCA
15851 TGTATGCACT ATATATTTTT TTCTGAGCCA CATAAAGAGT AAAACGCAGA
15901 CATGACGTGC TTTTACTCCT AAATACTTCA GTGTGTGTAT TCCCTCAAGA
15951 AAGGGCATTT TCTTCTGTAT AGCTACCGTA CACTTCTACA CTTTTCAAAA
16001 TCAGAACATT TACATTGATA CCATACTATG ACATGATCTG CAGACCATTT
16051 TCCAATATGC CAGTTGTCCC ACTGTGTCTT TTAGTACAAA AGAAAAAGT
16101 TTTTTTTCCT GGTCTAGGAG CTAATCCTGG AGCACATGTT ACATCCTGTT
16151 GTTTTAATCT AGAACCCTTC CTCAGTTCTT TATCTTTCAT AACCTTGACA
16201 TTTTGGGAGA GTACAATCCA TATATTTTGC AGAATTTCCC TTAGTTTGGG
16251 TGTGTCTGGT TTTTCTTAT AAGATTCATT TTATGCATTG CTGGCCAGAG
16301 TACCACAGAA GTACTGTATA TCTTACCAGA AAGCCTAAGT GGCATTTGCA
16351 TTTTCTAAAT GATCAATTTT AATATTATAT GGAAAGCAGA GTCAGAGATT
16401 CTCACATATG TCAAGATATT ATAAGTATTC CTGTTATATT TATTCTCCAA
16451 TTGCTTTTTC TCAAGAAAAT TTGTGGCCTT TCAGCTAGCT TTTCAAAGTG
16501 GAAGTTACTA CATAACATTA GGATGGGAGG GGTGGGGAAG AGCTTTATTA
16551 AAGCTTTAAG ATTGAGCTTT TGAGTATGTG TTGTATGTAA ATGAAAGTGG
16601 GCATTGATGC AGGGATTGGG CCTTTAAACC TTTGGCCAAG AATGGTATCA
16651 ATTATTATTA TTATTATTTT TTGGAGTACT TCTGCTAAAA CACTGAAATC
16701 AGTGTGCCAC TCTCCTTTTA GAAGTTTAC ACCTTTCCAA GGTACACTTT
16751 TTTTTTTTGA GACGAGTTTT GCTCTGTGCG CCAGGCTGGA GTGCATTGGC
16801 GCAATCACAG CCCACTTCAG CCTCTGTTTC CCAGACTCCA GCAGTCCTTC
16851 CACTTCAGCC TCCCGAGTAG CTGGGATTAC AGGTGCACAC CACCATGCCC
16901 ACTAGTTTT TGTAGAGATG GGGTTTTGCG CATGTTGCCC AGGCTGGTCT
16951 CCAACTCCTG CGCTCAATCT ATCCGTCTTC CTCAGCCTGC CAAAGTACTG
17001 GGATTACAGG CGTGGGCCAC CACTCCCGGC TTCCAAGGCA GGCATTTAAA
17051 TGTAATAAAT AGGGAGATAA GCAAGAACCC TGTGGACCT GGTAGAAGCA
17101 AACATTTATT AGTACTATTA CGTTGTTTAA AATATTAGCG CCTTCTATAT
17151 TCATGTCTTC CCAGAATTAT CAAAAACCT ACTCTATAGT TTATTTGGCT
17201 TATATCTCAG GAGTAATAAA ATTAGTTAAT AGTATTGGCA TCGTGGTTCT
17251 TTGTGTATTC CTCCCTTATC CCACCCCAAG TTGATTTTAC ATGATCTCTT
17301 GATCTAGTCT AAGAATGTTT ATAGTGATTA CGAGAAGTTC AGATTCTGGC
17351 TTTAACATAT ATAATTGTTT TTTAATCTGT AAACCAAAGA GAATGAGTTT
17401 GTTTAAACTA GAAAGATGGC AAGAGTAGTC TGGGAATTTT GTTCCATTCC
17451 TTAAAAGTCC TATAATAAAA TAAACATATC TTGTGTTTTA TTTTACAAT
17501 TTTTTTAAAC ATTAGTACAG AGTGCCACTT CTTATATTCT ATATCAAATA
17551 ATGAGCTACA TTTTCAATAA TAACCTCTGA GTAATTTTTG GCATTAAAA
17601 GCTGCATTAC AAAATAATTT GAGGATATAA TTTATAATCA CTTATGTCAA
17651 AATCACCTAT TTGAAATTAT GTATGAGGTT TTCAAAGTTT ATAGTGCTTT
17701 GGAAAAAATT TAAATGTTTC TTTGTTTATG TATCTTTATT ATAAGCTGTA
17751 GCATATATCA TGTAGTTGTC AAGGATGCTG ATAGATACTT AATATTTAAA
17801 GGAGACTTGT CTAAAGTTAG CTGTCCAGGA CTAGAATCTG GGCTTTTGG
17851 TAACAGCTCA TTGCTCTATT TACTTAAATG ATGATTGGAT TCGTTAGAAT
17901 TTCTCTATTT CTATAGCTGT CTATAGGTTT CTATGAAAAT ACTGTGTGTG
17951 TGCTTATACA TATATGTATA CCTGTAAGTA CAAAGTAGAA AATGAAAGTT
18001 CATTCTCTGC TTTTGACAAT TGTAATCCCC AGAGATAACC GTTATTAATA
18051 TGTTGTCTCA TGTTTGGTCA TACTGTTTTT TCTGTATTCT GTGTATTACT
18101 GTATAAATTT TACACAGTAA TTTGCATATT AAAAATGCTG GTCTACACCT
18151 GGCCCTTTTT TAAAACTGC AATTTATTAT GGCCAATTTT TTATACCAGT
18201 ATATATTGAT CAACCTTATT CTTTTAACT GCTGCATTTT ATTCATTACC
18251 AATAGATGAG ACATTTCAT TGGTTTGAAT TTTTCAGTAT TACAGATAAT
18301 GGTTCAATTA AATATTTAAG CTTTGTGCA CTTGTAGAAT TAATTCCTAG
18351 ACATAGAACC CTTATATTTT GATAGGTATT TCCAAATTTT TTCCCAAAT
18401 GTTTGTATCT CTTTACTTCC ACTCTCAGGT CTAATAATTT TCACTTGGAT
18451 TATCATATTT CTTACCCAGC CTGTTTTTTA CACTCTAAAC TCTTTTCTT
18501 TTCTTTTTTT TTTTGAGACA GCATCTTGCT CTTGGCCCCG TTGAAATGCA
18551 GTGGCAGCAG GACCAACCTG GGCTCAAGCA ATTCTCTCAA CTTAGCCTAC
18601 TGAGTAGCTG GGACTACAGA CACATATCAC CATGCCCAGC ATTTTTTTTT
18651 TTTTTTTTTT GGATTTTTAG TAGAGATGAG GTTTTGCCAT GTTGCCCAAG
18701 CTGGTCTCAA ATTCCTGAGC TCAAGCAATC CACCATCTC AGCCTCCCAA
18751 AATGCTGGA TTACAAGCGT GAGCCACTGC ACCTGGCCCA AAAGCTCTTT
18801 TTCTAATAGC AATATAAATT GTCTTTTACA GACTATACTC ATATATGTTT
18851 CTTCTTTCAG AAATAGGTGT TAAGTGTATC TAACATGGAA TGTATAGCTA

FIGURE 3, page 6 of 27

18901	TAATTCTCAT	TGTGAAACCA	TAGCCTAATT	TATTCATAT	TACAATTTAA
18951	AATTCATATT	TTTTAGGAAG	TTTTCTTAGA	TTAATCCGCC	TAGTTCAGG
19001	TGCTACAGTC	CCAAGATTTC	TTTCTTTT	ACAAATTTAA	TATAGGTAAC
19051	ATGACTAGAA	TTGTAGTCAA	AGAATATTGG	AACCTTGGAA	CTTCAGTATT
19101	TGAACTTTAT	TTTGAAATAT	AATTTGTTAT	ATTATAAAAA	TATTATAATA
19151	TATTGCACCT	GGAAGTTAGG	GGCAGTTTTT	TTTAATTCTC	TTTGTATCTG
19201	CTACACTGTA	AAGTGTCTATT	TATGTAAAAA	ATTCTTAATA	GAAGCTTCA
19251	GTTGTAAAGT	CTGCTGTACA	GACTTTAGAT	CAGGGATTGG	CAAACATATGA
19301	GCCATGTGCC	AAATCCTGCC	CTTCACCTGT	TTTGTAATA	AAGTTTTATC
19351	AGAACACATT	CAGACTCATT	CATGAACATA	TTGTCTATGA	TTTATTTTCT
19401	GCTACTAGG	CAGAATTGAG	TTGTTGCAAC	TGTGTGGCAT	CCAAAGCCTA
19451	AAATATTTAC	TCTCCTGGCT	CTTTGCCAAC	CCGTTT	TTATGAGCAC
19501	TTTGGCATT	TTATGTTTTT	GTTTTCTTTC	TATAGCACAC	AGTAAGATGT
19551	TCTGCCACA	TTGTGCATAA	TTTATGGGTT	TATTCAAGGA	TTTATGCAAG
19601	TGTAGCTGCA	AGAAAAAAC	CTAGAAGTGA	ACTTGCTAGG	TTGAAGAGCA
19651	TCTGTGTATG	TTAAATTTTG	TTAGCTTTCG	CCTTCCCAA	GGGATTATTC
19701	CATTTCATAC	TTAAACTACT	AATTTTGTGA	TAGGACTTCT	TTCTCCATAG
19751	CTTTGCTAAA	TTAATGCATT	CACACACTTC	ATCTTTACTA	ATCTGATAGA
19801	GGGAAATGAT	ATTGTGGATT	TGATTTGCAT	TTCTTTTTAT	GTGTTAGCTT
19851	GAGCTTATTT	TCATATTTAA	AAGCCAATTG	TATTTCTTTT	TCTTGAGCTA
19901	TCTTTTAAATG	TCCTTCCTGA	TACATTCTTG	AAGTCTGTGA	TACTCATATA
19951	AGATATATGG	TGAACATGTG	TCAAAGATTT	ATTTGACTCT	AATGAGGGAA
20001	CCCGCTGAT	GACAAGGCTG	ATTGAGAAGA	GGATGTGTGA	GATGAAGTGT
20051	ATATCATCAG	TGAAAGAAAG	CAAAATCTTA	CAGGGCAAAA	ACAAAACCAC
20101	AACTCTAAGG	GTTATGTGTTT	CTACTGGACA	GAATTCATTT	GCATTTTACC
20151	AGATAAAAAAT	TACTATTTTC	AATTTATCTT	TTACAAATCA	TTTTCTAATT
20201	TTACAGAGTC	TATTCCCTAA	TCAGTAGTAA	ATAGTCTTCA	AAATCTCCG
20251	CAGCGTCAGG	TGACTATTAT	GCAGGCTAAT	TGTTGACACT	CGGGCTTGAC
20301	TTTAAGAGAA	CATGCCATAA	TCTTTTGGCC	TTACTTCCAA	GTTTTGGATA
20351	ATTTTCTCTA	ACACATTTT	CTCTAATTGC	AATGATTTCA	AGTGATATTA
20401	TTTTCTTTTTT	TTAAATTTTT	TTACTATTTA	TTGATCACTC	TTGGGTGTTT
20451	CTCGGAGAGG	GGGATTTGGC	AGGGTCATAG	GACAATAGTG	GAGGGAAGGT
20501	CAGCAGATAA	ACATGTGAAC	AAAGGTCTCT	GGTTTCTCTA	GGCAGAGGAC
20551	CCTGCGGCCT	TCCACAGTGT	TTGTGTCCCT	GGGTACTTGA	GATTAGGGAG
20601	TGGTGATGAC	TCTTAATGAG	CATGCTGCCT	TCAAGCATCT	GTTTAACAAA
20651	GCACATCTTG	CACCGCCCTT	AATCCCTTTA	ACCTTGAGTT	GACATAGCAC
20701	ATGTTTTCAG	GAGCAGGGG	TTGGGGGTAA	GGTTATGGAT	TAACAGCATC
20751	CCAAGGCAGA	AGAATTTTTT	TTAGTACAGA	ACAAAATGGA	GTCTCCTGTG
20801	TCTACTTCTT	TCTACACAGA	CACAGTAACA	ATCTGATCTC	TCTTTTCCCC
20851	ATATTTCCCC	TTTTCTATTT	GACAAAACCTG	CCATCCTCAC	CATGGCCCCGT
20901	TCTCAATGAG	CTGTTGGGTA	CACCTCCCAG	ACAGGGTGGC	GGCCAGGCAG
20951	AGGGGCTCCT	CACTTCCCAC	ACTGGGCGGC	CGGGCGGAGG	CGCCCCCAC
21001	CTCCCAGACG	GGGCGGCTGC	CGGGCGGGG	CGCCCCCAC	CTCCCAGACT
21051	GGGTGGCCGG	GCGGAGACGC	TCCTCACTTC	CCAGATGGGG	CGGCTGCCGG
21101	GCGGAGGGGC	TCCTCACTTC	TCAGATGGGG	TCGCGGCTGG	GCAGAGGTGC
21151	TCCTCACCTC	CCAGACAGGG	TGGCGGCTGG	GCAGAGACGC	TCCTCACCTC
21201	CCAGACGGGG	CAGCCGGGCA	GAGGCGCTCC	TCACATCCCA	GAGGGGGCGG
21251	CCGGGCAGAG	GCGCTCCCA	CGTCCCAGAC	GATGGGCGGC	CGGGCAGAGA
21301	CGCTCCTCAC	TTCTTAGACG	GGATGGCGGC	GGGGAAGAGG	CGCTCCTCAC
21351	TTCTTAGATG	GGATGGCGGC	CGGGAAGAGG	TGCTCCTCAC	TTCTTAGACT
21401	GGGCGGCCGG	GCAGAGGGGC	TTCTCACATC	CCAGACGATG	GGCAGTCAGG
21451	CAGAGACGCT	CCTCACTTCC	TAGTACAGGG	TGGCGGCCGG	GCAGAGGCTG
21501	CAATCTCAGC	ACTTCGGGAG	GCCAAGGCAG	GTGGCTGGGA	GGTGGGGGTT
21551	GTAGCGAGCC	GAGATCACGC	CACTGCACTC	CAGCCTGGGC	AACATTGAGC
21601	ACTGAGTGAG	CGAGACTCCG	TCTGCAATCC	CGGCACCTCG	GGAGGCCGAG
21651	GCGGGCAGAT	CACCTGAGGT	CAGGAGCTGG	AGACCAGCCC	GGCCAACATG
21701	GCGAAAACCC	GTCTCCACCA	AAAAACACAA	AAACCAGTCA	GGCGTGGCGG
21751	CGCGTGCCCTG	CAATCCAGG	CACTCGGCAG	GCTGAGGCAG	GAGAATCAGG
21801	CAGGAAGGTT	GCAGTGAGCC	GAGATCGCGG	CAGTACAGTC	CAGCCTCGGC
21851	AACAGAGGGA	GACCGTGGA	AGTGGGAGAC	GGAGACGAGG	GAGAGGGGGA
21901	GACCGTGGA	AGCGGAGGT	GGAGACGAGG	GAGAGGGAGA	GGGATTATTT
21951	CTGTATGACT	TAATAATGAA	TTTCTAAGAG	GTCACTTAGC	TCACTGTTGT
22001	CTCTTCTAAA	ACATACTCAT	CTTCTCTTTT	CTCTTCTGTA	GGAACCTCATT

FIGURE 3, page 7 of 27

22051	ATACAATGAC	AAATGGAGGC	AGCATTAACA	GTTCTACACA	TTTACTGGAT
22101	CTTTTGGATG	AACCAATTCC	AGGTGTTGGT	ACATATGATG	ATTTCCATAC
22151	TATTGATTGG	GTGCGAGAAA	AATGTAAAGA	CAGAGAAAGG	CATAGACGGG
22201	TAAGTGTTTT	TAGTAAAAAT	TTTTAAAAAC	ATAGTGCATA	ATTAGATCTT
22251	TTAATAATAT	ATTTCTGCCA	ATGATCTCAG	GCTGCCAAAT	GTTTACATTT
22301	AATATAAGTA	AATGTCTACA	TTTCATATGT	GGTACATGTT	TTTTTCTTTT
22351	TCTATGTTTA	ATTTTTTTAG	TTTACTTATA	CCCTGTAAct	TTCCAGAAAG
22401	GATTTTCAGT	AGCTAAAAAA	CAAAGAAATA	CAATAAGAAG	ACAAAATAAG
22451	AAGGAAAGGG	AAAAATACAG	CACAGGAGTT	GGGGGGAAGA	ACAAAGCCAAG
22501	TTCCAGATAT	GGAGGTCAGC	ATGATTTTGG	GCTTTGAGCA	GCCCACCAGC
22551	TAAGGCAAAA	AAGGAAACTC	ATTGCATAGC	TCTTACCTAT	GGAAAAAGAA
22601	GAAATCTACT	GGGGGCGAGT	GGTCTGTGGG	GATTTTGCTG	TTTTCTTTTA
22651	TCTCCTTTCC	CAGCATTTGA	TTCTGAGATA	TTTCTCAATT	TGGCTCCCAA
22701	ATAAAGCTTA	TTGAGTGTG	TAATGGTTTA	CTGTTTTTTT	TAAAAATGGC
22751	TTTAACATAT	AAAAGTACAA	CTTATGGATC	CTTTTTGTTT	GTGGTCGTGA
22801	CTTACTGATA	ATATAATCCA	AAATACATTT	TTTATTTTGT	ATTTATTTAT
22851	TTATTTTTGA	GACGGAGTCT	CAGTCTTCTG	CCCATGCTGG	AGTATAGTGG
22901	TGTGATATTG	GCTCACTGCA	CCCTCCGCCT	CCTGGATTCA	AGCGATGCTC
22951	CTGCCTCAGC	CTCCTGAGTA	GCTGAGACTA	CAAACGTACG	CCACCATGCC
23001	TGGCTAGTTT	TTATACAAAA	TACGTTTTTT	AAAAACAAT	TTTTTTTTTG
23051	GAGGTCGGGG	GACTGTCGCC	CATTCTGTTG	CCCAAActGG	AGTGCAGTGG
23101	TGCAATCTTG	GCTCACTGCA	ACCTCTGCCT	CCCAGTTTCA	AGCGATTCTT
23151	GTAActCAGCC	TCCTGAGTAG	CTGGAATTAT	AGGTGTGTGC	CATCATGCCA
23201	AGCTAAATTT	TGTATTTTTA	GTAGAGATGA	AGTTTCGCCA	TGTTGGCGAG
23251	GCTAGTCTCA	GACTCCTGGC	CTCAAGTGAT	TGGCTGACCT	CAGCCTCCCA
23301	AAGTAGAAAA	TCTTCTTGAA	AAATAAAATT	CCAAATCTCA	AAAGGCCCTA
23351	TATAATTTTG	GTGTTGAAAA	TTTACTTGTC	AATGAAAATG	ACTATTTACA
23401	CAAATTATAA	GCTTCCATAT	TAATATATAT	GTGTGTGAAC	CTGAAATTCA
23451	AATTTTATTA	TATTGTTTAT	GAAAGGTACA	GCCTCTGAGA	TTCATCAGAT
23501	GGTATTTACC	TTTAGGGCAT	ATCTAAAAAT	AAAATACAGT	ACATGAAATC
23551	CAGTGCCTTA	ATCCAGTGAT	TCTTAAACTT	TTTGCTCTCA	GATCCCCTTT
23601	AAACTCTTAA	AAGATATTGA	AGAGCTCCAA	GGAGGCTTTG	TTTACGTGGT
23651	TTTTATCAAT	GGATATTTAC	CATATTAGAC	ACTGAAActG	AGGATTTTAA
23701	AAAAAAATAA	TTCATTAAAA	AATAACAGTA	ACAAAACCCA	TTACATGTTG
23751	ACATAAAATA	CATTTTACG	AAACTATATT	TTCAAAAATT	AGTGAGAGAA
23801	TGACATTGTG	CTACATTTGT	TATAAACTC	ATTATTGTCT	GGCTTAATAA
23851	AACACTGCTG	GATTCTCATA	TCTGCTTTTG	NNNNNNNNNN	NNNNNNNNNN
23901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
23951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
24851	NNNNNNNNNN	NNNNNNNAAA	TATTGATTCA	CTGATTATG	TGGATCTTTT
24901	AAATGTTGAC	ACTTATATAA	TATAATACAA	TATTTTAAAA	ATCACATTTG
24951	TTAATTTTAC	CTTTGATCTA	TTCAGAAAAG	ACTCTAAGTA	TTGGGAACCT
25001	ATCATCCTCA	CAGTGATAGA	TACAAGTTTC	CTAAAATTCT	GATTTTACT
25051	GGAGAGCTCA	AATTCTATCA	TTGGAAACAA	ATACACATTT	ATTTAActTTA
25101	AAAATGACAG	GATTACTTGG	TTTCATTATT	GAGAAAATAC	CTGTCAAATT
25151	CCCAAGTCTG	GAAAACCATG	GTTTGATGTC	ACTCTTTCAA	GTAAAAATGG

FIGURE 3, page 8 of 27

25201 CATTCCATGT AAGAAGTGTC TAGTTTATTA TGCAACTCAA ATAAATTACG
25251 CAAGTGCTTT TCTTTAGGAC ATAACCTCAT ACATACTTCC ACAAGCAGCA
25301 GATGTGTGTA GTTATGCATA GTTCCTTATG CATGGTTCTT ATTTTCATCAC
25351 ACAAATATTT AAAAAGACTC AGTGATTGAG ACGTAGCAGT TTTTACTGCT
25401 TCATCAAAGA TGCTCTTATT TGAAACTGGC ATAATATGAT TTATTTATTT
25451 GATTTTACTG GGAAGCATGG CAGTCAAGAA TGTAAAGACT GCCAGTACAT
25501 TTGAGTGCCA CTGCTTGATT TTTGCTATGG AGTCAGCAAT TTTGCCACTG
25551 GTTTTGCAAT TTCAGTAAAA ATGTCAACAC AGTGAAAAAG GCACATAATG
25601 TCTTGTATTA TTTTGTAAAC AGTTTATCT TGCAGACCCC TTGAAAAGGT
25651 CTCGGGGATC CTCCAAGGTG CCAGTAGACC GTACTTTGAA AATCACTATT
25701 TTAATCCAAA GTGCCTAGAT CAGACACACT ATAAATCCTG TGTCTTGTAT
25751 GATCATTAGG TGAATACATT TGTACTTAGA AGTATACATT CAGAGACATT
25801 AACAGTATTC AGGTGGGAT TTAAGTATAT TTTAAAGTGT GGTACCTAGA
25851 GAGTATCCAT GACACTATGT TCATAAAATT TTAGAGAAAA CTGAGATCAA
25901 AGGAAACCAA AACAGGCTGG TCATAGTGGC TCATGCCTGT AATCCCAGTG
25951 CTTTGGAAGG TTGAGGCAGA GGATCGCTGG ATCCCAGGAG TTTGAGACCA
26001 GCCTGGGCAA ATATGGAGAC TATCTCTACA CAACAAAACA AAAATTAGCT
26051 GGGTATAGTG TCTTGGCCTT ATAGTCCTAG CTACTCGGAA AGCTGAGGTG
26101 GGAGGATCCC TTGAGCCTGG AAGTTCTAAG TTACAGTGAA TTATGATTGC
26151 ACCACTGCCC TCCAACCTGG GTGAAACAGC AAGACCCTGT CACCCTCCAA
26201 AACAAACAAA AAACACTTTT TTCTCTGAGT ATGTAAATGG TTAGTGACA
26251 GTCCTTGAAA ACATTGCAAA TAGTATAGCA ATATATGAAG TAGCCAGTAT
26301 GTGCTCAGC TAATTTTATC AATCATCTCT TCCTAGACCA ATCAAATATT
26351 TTTCAATATT TTGATCCATG CTTATATGAA CAAGATTTT TAAAGCTGGA
26401 AAATTCACA CATTATATA CTTACTATTG TTCTTAAAA TAATTTTTTT
26451 TTTTTTTTTT TAAGCAGAGT CTTGCTCTTT TGCCCAGGCT GAAGTTCAGT
26501 GGGGCGATCT CGACTCCCTG CAACCTCTGC CTTCCAGGCT CAAGCAGCTC
26551 TCGTGCTTCA GCACCCCAAG TAAGTGGGAT TACAGGCATA CGCCACCACA
26601 CTGGCTAATT TTTGTAGTTT AAGTAGAGAT GTGGTTTCGC CATGTTGGCC
26651 AGGCTGGTCT CAAACTCCCG GCCTCAAGTG ATCCACCTGT CTCAGCCTCC
26701 CAAATAGTTG GAGATTACAG TGGGAGCCAC TGCGCCCGG CTACATTAAA
26751 TTTTAAAGCC TTTCTATGTC AGTGCAATA CCCAACCTAA TTCTTTTTTT
26801 CCGTGAACCT TTTTGTATG CTTGTAGCCT TCCTACCCCA GATTATTTTCG
26851 AAGCAAATG TCATTCTGTA ATTTCAAATA TTACTATTTT AGTATTTTAC
26901 AAAATGGTTG CAGTTTAATT GTTGTCCCTT TTTTATTTAT TAGCTTGCAT
26951 ATTTCTATAG AGAGTTTACC CCACATCAAC CATTGGGATT ACCTGAAGTA
27001 AGGCTGGTAT AGGAAAGGGA GAAATCTTGA AATACTAGT TCCTTAGCAT
27051 CCTCAAAGTT GACCAATGAG ATTTTGTGCT TGTTTGGTTG TTTTCTCTG
27101 TGTCTTCTGG ACTCATGGAT TTAAGTATAT TTGTGGTTTA ATCATCACTG
27151 TTATTATCTT TATTGATGTT CATGTTATTT TAGATTAGTG GGAGCTTTTT
27201 TAGTTTGCTA TCTGTGCTCT TCGTCATGTC CTTAGATAAT CCTAATCCTA
27251 ATCCTGATTC ATCGTAGACA TTTCCCGCAG CAAACCTGGA ATCAGCCATT
27301 TCTCAAGGAG CTCTCTGATT CCATTGAAGG AAAATATAAT ATAGGTACAA
27351 TCTAGGCACT AGGTGATACT TGTACTTCT GGGTTGGCTA TTGTTTCTAG
27401 CCTCCTAAGT TTATATGACT GTACTAATTT GAATTCATAA CTATGGGACT
27451 AAACCTCTAA TTCTTAAATC TGCATTTCTT TTAAGTCATG CCAAAAATCT
27501 GAACATCACA AACATAGTCA TTTCGTTTAC CCCACAATAC ACACATACAA
27551 CATTGTCAGT ATAACAGTAC CAACACCATC TCCAACAATA TGCTACTGA
27601 AAAATTTTAG GTAATCTGTC TCCAGCCTCC CAGGTAGCTG GGACTGCAGG
27651 TGCACACCAC CATGCCCTGG TAATTTTTTT TTTTTTTTTT TTTTTAAGA
27701 GACTGGGTCC TTGCTATGTT ACTCAGGCTG GTCTGAAATT TCTGGCCTCT
27751 AACAGTCCCT CTGCCCTTGC CTTCCAAAGT GCAGAGATTA CAGACCTGAG
27801 CCACCACGTC TGGCCTATCC TTTATTTATT CCACCAAAGT TATTTATACA
27851 AATTACTTTG TTGTAAAGTC CCTTGAATA GTTCTTCTG TGGCATTATG
27901 TTACCAGTTA GATGCACCTT TGATTCAATTT AACTTTACTT CAATTTTTTA
27951 GGTTTGCTTT TTAGATTTAG TTTTGTTTTA TTATACATAT ATGAAGTATT
28001 TCCACGGTTC CAAAGTTAAA TGAACAAAAC AGGCATGTTT AAAGAAGTCT
28051 AGTTTCTATC TCTGTCCCAT CCAACCCATT GTCTTCTTCC CCTATAAGT
28101 AATAATTTAC ATTTTAAACT TGTGGTTTAT CTTCTGATTT TTAATAATAT
28151 AAGCATAAAT ATTTATATTC CTGTCTTTTA GCATGCTTTT AGCCATCTTG
28201 CTTTTTCTCT GTATAATGCT AAATATATCT CATTCTTTTT AATTGCTGCA
28251 GAATTTCTCA TTACATAGGT ATACTGCAAT TTATTTATCT GATGCTATGT
28301 TGATGAACAT TTAAATGATT TCCAGATTTT AGGAACGGTG ATGATTGAAC

FIGURE 3, page 9 of 27

28351	TCTCTGTACA	TATATCTTTT	TTACTTGGTA	CACTCCATCA	AGCAACTACT
28401	TAAGTGACTG	ACTATGATGC	TGTGCAAGCA	GTTATATAAA	GAAAACAGCA
28451	GTGACTCAGC	CTGAAAACGG	CTTAATATTA	TCATGTTTTC	TTACACATTA
28501	TTTTTATTGA	GGAAAAGCAA	CATGGAGTTT	AGTGATTATT	TTTGAAAGAA
28551	ATAACCTATT	TCTAATTCTA	AAGAATGGTT	ANNNNNNNNN	NNNNNNNNNN
28601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
28951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
29001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
29051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
29101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNGAGTCTAG
29151	CTCTGTCCAC	CAGGCTGGAG	TGCAGTGGCA	CGATCTCTGC	TCAGTGCCAC
29201	CTCCGCCCTCC	CGGTTTCAAG	TGATTCTCCT	GCCTCAGCTT	CCCAAGTAGC
29251	TGGGATTACA	GGCGTTCGCC	ACCACACCCA	GCTAATTTCT	GTATTTTGTAG
29301	TAGAGAAGGG	GTTTCACTGT	GTTGGCCAGA	CTGGTCTTGA	ACTTCTGACC
29351	TCGTGATCCA	CCTGCTTCGG	ACTCCCAAAG	TGCTGGGATT	ACAAGCGTGA
29401	GCCACCACAC	CTGGCCAAAA	ATATGGGTTT	CTAAAGCAAC	AGTCCTAGTA
29451	CAACAGAAGA	GAGGTGTTGA	CTAGTTAGGG	ATTTAGGTTT	AGAAGTACAT
29501	TCTTAGTAAG	AGAGGTGAGA	CTTACCTTCT	TGTGTTTTAG	TATAGTGAGA
29551	TCTGGATCAA	ATCTATTACT	CTTATTAATC	TCCTAACTTC	CTACACTATA
29601	TCCAGTAGAG	GACACTTTTG	CCTTACACAG	TAAAGAAAAG	GCCTCTGGAC
29651	TCTACCAATG	GGATCGGAGC	TCTCCAAACC	TGCATATTAA	AAGGCCTATA
29701	AGTTTTGGGG	GGTCCCTTTG	TCCACATGAT	TATTCTGTAA	TACATTGTAT
29751	TTATGGACAT	GGTATTATTA	TACACAGATC	CTGTCTTTTA	AAGAACATTA
29801	TAATCCACTT	AACTGCCTAG	ACCAGAGAAT	GACCGATAAT	TCAAACCATA
29851	TTGTCTTACA	GAAGACATAT	ATAAAAGATG	GTTATGTGTA	CCAATTGAGG
29901	TTCAAATTTG	ATTCAATTTA	AAACAATCTA	GGCCAGATTT	TATATAGTTT
29951	GTGGACCCTT	TGCACTCAAA	TCTCAAGGTT	CTTATTAAAA	TGCAGATCTT
30001	GGCTGGGCAC	GGTGGCTCAC	ACCTGTAATC	CCAGCACTTT	GGGAGCCCAA
30051	GGCAGGTAGA	TCATTTGAGC	TCAGAAAGTC	AAGACCAGTC	TGGCCAACAT
30101	AGCGAGGCCC	AGTCTCATTG	AAAGAAAAAA	AATTTTTTAA	TAAAAAATAA
30151	AAGCAGATCT	TGGGTAAAGA	CATGTAGTCT	GGTTTACAGG	TATTAACAAC
30201	TGTCTGTAAT	GTAGTGATTT	TGCTCCAGAC	TTACCTTTTC	CATTATTTAG
30251	TTCTGAAATT	ACTGTTCTAT	GTATGGTAAA	TGAGAAAAAT	TGCTAGATTC
30301	TAGAACTGTG	GCTTCTATTC	ATAGTTGGAA	AAATGAAGCA	TAAACATTTT
30351	TAATTTTCAGA	TCAACAGCAA	AAAGAAAGAA	TCAGCATGGG	AAATGACAAA
30401	AAGTTTGTAT	GATGCGTGGT	CAGGATGGCT	AGTAGTAACA	CTAACAGGAT
30451	TGGCATCAGG	TAAAGAAAAA	TTTTCAAGCA	ATCCTTTTTT	AGTTAACAGA
30501	AGTATAAACT	GTTCTTCCCT	CCTTCCCTCA	ATTTTTTTTC	AGGTACCATT
30551	GGATTTTAAA	AAGCATTTGT	TTCTCTTCTT	CAAAAAATCT	CCTTAAATAT
30601	AAGACTAGGA	GGCAGAGGCT	TCCAAGTCTA	GTCTTGGCTC	TATCACTTTA
30651	CGTGTTTATC	CAGCTTGGTT	GATCTTTCTG	GACTCAGTTT	CTATATCTGT
30701	AAAATAAGTG	GTTTGGATCA	GATGATCAAT	AAAGTATCTT	TTGATATTAA
30751	CATCGTAATA	AATAGCTAAT	ATTTCTTGAG	TGCTTCCTAT	GTNNNNNNNN
30801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNTGGAAGAT
30851	TATGTTTCAGA	AGACCATAAA	AATTAATAAT	TTTGTGGAGA	ATAAAGTACT
30901	GATAATTCTA	ATTGGCATGC	ATAGTAATTT	TATGGCCTCT	GTGTATGTAA
30951	CCCACTGATC	TCTTTATGTA	AGAAGGACCC	AGATTTGACC	ATAAATTTGT
31001	GTATTTTTTTA	TATTCTCACA	ATAAAATAAT	CTTGATATAT	GGTTTTCTGT
31051	AATTTAAGAA	AATATTATTC	CTATGAGTTT	CAATAATTAT	TTCTAATGGA
31101	CATTAAATTT	TAATGAAATT	GACATCATTT	ATAAGTCTGT	TAATTAAGTT
31151	ATCGATTGAA	AATTAGATTT	GTGAACCTCC	TGCCAAGTAG	CTGTCTTTTG
31201	AAGATATTTT	AGTATCTTTT	AAACATTGTT	TTTCAGATCA	CAATTAATTT
31251	GAATGATGTA	ACTTTTTAAA	ATTCCAAACA	AAAATAGCAC	TTTTATTGTA
31301	AAAATAAACT	CTTTACAGTT	TATAACTAAA	ATTTGAAAAT	CTTAAATTTA
31351	TATGTAGTTC	ATAAATGACC	CTTTATTTAG	GAGTCTCCTG	CTTTCTACTT
31401	GCCTTTTAAAC	TAGATTGTTC	TCGACTCCCA	AAAAATTGAC	TTAATTTTTT
31451	TACCATCTCC	AACATGTTTT	TATAGGGGCA	CTGGCCGGAT	TAATAGACAT

FIGURE 3, page 10 of 27

31501 TGCTGCCGAT TGGATGACTG ACCTAAAGGA GGGCATTTCG CTTAGTGCGT
31551 TGTGGTACAA CCACGAACAG TGCTGTTGGG GATCTAATGA AACACATTT
31601 GAAGAGAGGG ATAAATGTCC ACAGTGGAAA ACATGGGCAG AATTAATCAT
31651 AGGTCAAGCA GAGGTAAGTC TTGCTTTGTC TCAAGATGAA TTAATAATTG
31701 ATATAGCAAA ATGTTTCCAA TTCATTTAAT TATAGAACTA ATCACATATT
31751 AGATGATTAC ATACACATCA AATGGATCCA CCCTCAACAC ATTGCAGCAA
31801 GAAAGAATTA AGTGCAATAT TGTTTCAAGT AGCTTTTTTA TTAGTTAACT
31851 GCATAGTCAT ATAACAAATC CTCTGGATTG TGGTGCAAAT ATATTTGAGC
31901 TGTAGTAGAA AAGAAGTGAT AGTTATTGCA GTAAGATCTG TGTAAAGTTA
31951 CTAAGAAGTC AAGTTATTAA AACTAATATA TTAATAAGAT TGGGAAGTTT
32001 GAATTATGAA AGTATTATCA AATAATTAG TAAAATCAAC CTACGTAGAG
32051 ATACATTGAA GATAATCAGA CATTTTTATT TGTGGCATTG CAGCATTTAA
32101 ATGATTGATT TACTATGATC TACAAAGAAC ATTTTAGAAC TTAGGATGTT
32151 ACATGTATAT TTTTACATG ATGACATGGA TATATTTTTT AAATTTTGTT
32201 TTAGCTGAAC TTTAGAGCTA AAAGGTATAC ATTTGCGGTA AGATGAGTAG
32251 TATGCTGTTT CTCACCTGGC TTAATTGAAT TGAGTTAAT GATCTGAAA
32301 GTTGCAGCAG AATGAAATCT GAGTGGTGAT GCAATTGTT TCCACTGTTT
32351 CCAAAAAGTG GTTTGTAGGC AGAGATTGAA GTATAGCTGA GATGTGTTGG
32401 TAACAAGACT TTAGGGATTA GGAAAAGAT TAAATGTGCT CAGGGTTCCT
32451 TGGTATATGT AGGCATTAAT TTTTGGACTC TACTTAAATA TTTGTTCAT
32501 ATAAAGTTT TATTATGTG GAAATAAACC AGGAGACTTT TACACATTTT
32551 ACTGAAGTTT CTTTCTTTT TTTTTTTTTT TTTTTTTTTT TGGCCGGTGG
32601 GATGGAGTCT CACTCTGTT CCCAGGCTGG AGCGCAGTGG CACGATCTCG
32651 GCTCCCTGCA ACCTCCGCTT CTGGGGTTA AGCGATTCTT CTACCTCAGC
32701 CTCCCAGTA GCTGGTATTA CAGGCGTGCG CCACCATGCC CAGCTAATTT
32751 TTGTATTTTT AATAGCAACG GGGTTTCACC ACATTGGCCA AGCTAGTCTC
32801 GAACTCCTGA CCTCAGGTGA TCCACCCGCC TCAACCTCCC CAGTGCTGGG
32851 ATTACAGCG TGAGCCACCA TGCCTGGCCG TTTACTGAAG TTTCTTATGA
32901 CAAGCATTTG CATTAGAGGT GCAATGTAAA TTAAATTCAT ACTCTCGAAC
32951 TATTTTCTTT TTAGGGTCCT GGTCTTATA TCATGAACATA CATAATGTAC
33001 ATCTTCTGGG CTTTGAGTTT TGCCTTTCTT GCAGTTTCCC TGGTAAAGGT
33051 ATTTGCTCCA TATGCTGTG GCTCTGGAAT TCCAGAGGTA AGCCAAGTAA
33101 TATTTAGTGT CATTAAACAT TATTATGATG CTTATCTTTT TGACCTTAGT
33151 GATAATAAAA GTTGGCTTTT CTGGAGGGAG GGGATAGTTT GTTCATAATA
33201 TGAAAAAAA ATTTTTTTAA GTATAAGCTG ATGGTAGACA TCATTGAAAA
33251 ATATTGTTCC CCATAGTCAT TTGGTCATT ACTGTGAAGG CTGATTTTTT
33301 TTTTCTCTCA CCACATAATT AACACATGAC TAGGCAAAAT TTCAGACTAT
33351 TTAGTTAAAC ATCAAGAGCC TGGAAAGAGT ATCTGTGAC CTAATGTTCT
33401 TTGACGGGTT AGTTGTTACT TTGCTGTTAT GACCCTGAAT TTTTTTTTTT
33451 TGAGACTGAG TCTTGTGCTG TCGCCCAGAC TGGAGTGCAG TGGCGCAATC
33501 TCAGCTCACT GCAACCTCTG CGTCCCAGGC TCAAGCAATT CTGTGTCTC
33551 AGCCTCCTGA GGAGTTGCGA TTGCAGGCAC CTGTCACCAT GCCCTGCTAA
33601 TTTTGTGATT TTTTGTGTTG TTTTTTTTTT TTAGTAGAGA TGGGGTTTCA
33651 CCATGTTGGC CAGGCTGGTC TCAAACCTCT AACCTCAAGT GATCACCCGC
33701 CTCAGCCTCC CAAAGTGCTG GGATTACAGG TGTGAGCCAC CACACGTGGC
33751 TATGACCTG ATTTTGATTC ATTCACTTT TATAATTACC TTTTGATTAG
33801 ATAAGTTAAT TATTCTTGAA TTTGGCCATT TTATGCTTTG AGAAAGTAGT
33851 TAATCACAGT GGGTCAACAG TACAACTTT TGGGTTTTAT TTTTCATCAC
33901 AATAAAGTAG AGTTATACAT AGGATTGATT GAACTTGATT TGAACCTATC
33951 TCTTCTCTTT TATTTTCTG GAGTTAAATA AGTTACCAAC TTTTCTTAA
34001 TACATTCTT TTTAAATGG AATTGTATTG ATCCTTTAAG TTTGTATTAA
34051 GAATATCTTT CATAAAAAGC AATATCATGC AGTATATAAC AGTTGTTACT
34101 CATTCTTGAT ACATAAAAA CTATTGCACA TAATTACAGG ACCTCAGAGA
34151 AAACATAATA TTCTTATTT TAACATAATG GCCAAAATAT ATTTAAATA
34201 TTATGCTTAT TTTTACAACA GAAATATTCA AATTGCCCCT TTTTTTGGGT
34251 ATGTAATTAT AATCCTTATA ATTAAGGTCT GTATTCAATT TAACATGGCC
34301 TGATATTTTG ATTTTGGCCT GAGATAGTGT TGCCCTCTCT CCTTCTTGG
34351 GTAGAGAATT AGATTATAAT ATCAATTTAT TATATGTAGC ATAATAGGCA
34401 AGTTTTTCGAA AAATTAACGT TAAATTTTTC TGTAGACTGC TAAAATTTGC
34451 AAGGTTGTTT TTGTGCATAA AACAAGAAAA TAACTTGAT TCGTTACATT
34501 CTCATGTTTC TTAAGGACA TTAAGCTGCC TTAATCTTTG CCTGTAGAT
34551 TAAACTATT TTAAGTGGAT TCATCATCAG AGGTTACTTG GGAAATGGA
34601 CTTTAAATGAT TAAACCATC ACATTAGTCC TGGCTGTGGC ATCAGGTTT

FIGURE 3, page 11 of 27

34651 AGTTTAGGAA AAGAAGGTCC CCTGGTACAT GTTGCCTGTT GCTGCGGAAA
34701 TATCTTTTCC TACCTCTTTC CAAAGTATAG CACAAACGAA GCTAAAAAAA
34751 GGGAGGTAAG TGTCTTTTGT AGTTAATTTG ACTGAAAAAT ATATATTATA
34801 TAGTATTTAT TTAAGTAAAG AATTTCTTAG TGTAATAATA ATAAATTCTG
34851 TATTCAGATA AAAAATTTTG AGATTTGTGC TTCTGTTTTT CCTGAATAAT
34901 CTATAACATC TTTCTAGAAT CCATTCCCAG TGCTGCTCAG TTCGTCTTAC
34951 ATTTTAGAGA AGCTTTAGAT AGACAGCTGG TGTCATTGG GTTTCAGCTG
35001 CATTTACGA AGATCTTCCT GTTATCACTT TACCTTACAT CTTTCTCTT
35051 CTGAAGTGT TTCTAAGCTT AGCTTTGTTT TTCACTCTTA CTTTCAACAT
35101 TAAGAGGTTG GGAAATCTTA ATAGCTATGT TTTCTCTCTG GAGGCAGTGT
35151 CTGGTGCCAG TGTAAGTGGT GTGTGATATG AAAAATGCTA TCCAGTGCTA
35201 TGGGGAGGTT CTGAGGGCCT TTAGAAGCTC TTGAAGTTTA AATCAGAAAT
35251 TCACATTAAG GAGATTACAG GAAATCCTTT TCATTTGATT GTTTAAGGCA
35301 ATTTCTTTTA CCATTTCTTT AGGCCAGCCT GAGATCTTCT ACAAGACCTT
35351 GAAACCTTAT ATATATTATG GATTTCTCTT GATGTTTCCA TATTGCTCTG
35401 GGCATTTTCC TGAATCCTTT ATATTAGCTC TAGACTTTGG GAGCCCAGTC
35451 CCTTCTTATT TTCCAAATCT AAATCTACAG CCCTAGATGG TACAGAGATC
35501 TTTGAGTTTG TAAGATATGA TTTTGTGAAA AACATCTCAT TAAATACTGG
35551 CAGAACCTTT TCATCTTGTT GAGTTTTTTA ATGTACTGTA ACCAAAAAAG
35601 TAGAATATTT TATCAAACTG TTTAATCTTC AATTGAAATA ATTCTAGTAC
35651 ATTTTAAATG TCGCATTAAG ATATTGTCCT TGCATTGGAC GTAGATATCC
35701 CAAAGTGGA ATACTTCAGA TTGTCGTAGT TTCATCTCTG AATAATTGTG
35751 ATTCAGTAC TTTATAACAA AAATAGCTAG CATTATTGAT TACTTTCTGT
35801 GTATCTGGTA CTGTGGCAGA TACTTTACTT GGATTTTAA ACTTAATTTT
35851 ACAGTAATTT AGTAATATGG CCCTGTTATC CTCATTTAGT GATTAGTAAA
35901 CTAGGGCTGA AAACAGCTAA CTAAGTTGCC CGAGACTACA TACCTAGTAA
35951 GTGGTGGAAC GTAGGTTAAA ATTCATTTT CTTTGACTTC AAAGTCTGTG
36001 GTCTTACCTA CTTACATTAC TGCCCTTACG ACTATGTGGG TATATATTTG
36051 TGTGTGTTCA AAACAACTC AAAACCATCC TGTAGCGTAG CAAGTTAGTG
36101 GCTAAGATGA AGCTAGAGCA TTTGCCCTCT CAATTCAATT CCATTACTTT
36151 CTGTTGTACC TTTTATTTT TTGGTAAGAC TTTTACTTAT TCTAAGTTCA
36201 AAAAATGTAA TTTATTTAGT GTTTTGAGAAA TTAAGTTTAC CTAAATTTTA
36251 ATGTTTCATC TGTAAGTATT AGTTAATGTT TAATACGTTG TTATTCTGTC
36301 ACCTTAGTGT ATATATAAAT GGCAAGAATT CACGTTAGT TGAAAGCATT
36351 AAGGTCCCAT AGTTTTGTGT AGACAAGAGG GGAGAGCGTT GATATTTTTA
36401 AATTAAATGC TTCTTAGATA CGTATGAAAT GGATTAAAC ATGTATATGA
36451 GTTATAGATA CCTAGGTGTT AGTTTGGTTG TAAATTCAGG ATCAGGACAT
36501 TCAAATAAAT ATGTTTGCTT TCCTCTTAGT GGAGGAAAA AAAAAGAAGC
36551 TAAATTTGCT CCCTTCTCTC CCCAAATAAG CAGAGTCTAC ATTTAATGC
36601 CAACAATTTG ATTAACAAAC ATATTTATTT ATTTTAAAT CACCAAACTT
36651 TTATAAAGTA TTTACTGGTG CCAGGCACTG TTCTAAAGCA CTCTGTATAT
36701 ATTTACTCAG TCCTTAAGAG CTAAGTAATA TTATCACGTT TCCATTTTAG
36751 AGAAACTGA GGCACATATA GGTTAGGTTA TCTACCCATA GCCATACAGC
36801 TAGTAAGTAG CAGAGCCATG ATTTCAACAC AGCAGCCTGA CTATGGAGTT
36851 CATGATCTTA ACCATTACA GCTTAATTTT TATTATTTAT AATTTCTCTT
36901 CTGGAAATGT AACAATTGAC CATTTGAAGA AATACTTTAG GTAGCTTTGG
36951 ATATTTGCTG TATTAAAGTA GTGAAAGTAA TACAGACACT TGGCTGGGCG
37001 CGGTGGCTCA CGCCTATAAT CCCAGCATTT TGGTAGGTTG AGGCAGGCAG
37051 ATCACCTAAG GTCAGGAATT CGAGACCAGT GTTGCCAACA TGGTGAAACC
37101 CCGTCTCTAC TAAAAATACA AAAATTAGCC GGGCGTGGTG GCAGGCGCCT
37151 GTAATCCCCA GCTACTCGGG AGGCTGAGG AGGAGAATCA CTTGAACCCA
37201 GGAGGTGGAG GTTGCCAGTGA GCTGAGACGA CGCCATTGCA CTCCAGCCTG
37251 AGAAACAAGA GAGAACTCT GTCTCAAAAA AAATAAAGGA ATACAGACTC
37301 TTAGAAAAAT AATTACAAAT AAAACCCTAG TGAAATTATA GGTATAGTTA
37351 GGTATAGTTG GCTTACAGGT GGGAAAGTAG CCATTACCAA CTGATAGACT
37401 GGGGAGCTGG AGAGAGGACA CGGAAGAGTG TCCTTGGATT TTTCNNNNNN
37451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NAAAATTGTC TATATTCATT
37701 GCCTCTCCTT CTTTACACCC TATTCACATT AGTATATCTG GCAAAAATTT
37751 TTTTAACTG AATGGTAAAT GCATGACTGA CCTTCAATT AAAGCCAGGA

FIGURE 3, page 12 of 27

37801 GAAAGAAACA AATCTTAATA GAAGAAATGA ATAGTTACCC TTTGCTTAGG
37851 GAGCAAGGAA ACATGCAAGT TAAATTCAGA AAATCCATTT GGAAATTC A
37901 AGTAACATGA AGAATTTTTA TTTGGTATGT TTGAATTTCT ATGAAATTAT
37951 GAAATAAGCC ATATCCTCTT TCTAGGTGCT ATCAGCTGCC TCAGCTGCAG
38001 GGGTTTCTGT AGCTTTTGGT GCACCAATTG GAGGAGTTCT TTTTAGCCTG
38051 GAAGAGGTAG GTGAAAAGAA TACAACAATT AAAATTATAT ATAATTACCA
38101 TTACAAATAT ATTTTCACACA TTTTCAGTTT GTAGGTGATG TAATAGGTAG
38151 AGACTTTGTT TTCAAATTTA TTTTCTAAA GTTGTTTTCC ACTCATTCTT
38201 AATAAAAAGT AAATGTTATT CATGCTCCAT ACCTGGAGGA AACTTTTTAA
38251 AAATTTATTA ATGTATGAAT GTTAGTAATT ATTTAAAATC TAACTTTGTT
38301 GACATATTTA AAGTAAGAA GATGTGAATT TGACTTAATA GAGGACATGT
38351 GAAACAATCT ATTTCCATTG GCTAAATTCT GTATTTTAG TAGAGATGGA
38401 ATTTACCAT GTTGGCCAGG CTGTTTTTTT GTGGGGTTTT TTTGTTTTGT
38451 TTTGTTTTGT TTTTGT TTTT GAGACGGAGT TTCCTCTTG TTGCCAGGC
38501 TGGAGTGCAA TGGCGCGATC TTGGCTCACT GCAACCTCCG CCTCCAGGGT
38551 TCAAGTGATT CTCCTGCCCTC AGCCTCCCAA GTAGTTTTTG TTTAAAAAAT
38601 TTTAATCAAT TCCTATGTTG AGTTTTAAAG TTTTCCCCT GTGATTATTT
38651 CTGATACAGT TAGTGATGTT AAAGAAAATA ATTTTAGTGA CTTCAGTGGA
38701 TTATTTTGTT TTTGTTTCT TAATAGGTGT TTAAGACTTT TCTTTTACA
38751 TAAAAATGTA ACCAGGAATT TTTTTTAAT TTTTTTTGAC AAATAATAAT
38801 TGTTTTTGTT TATGGGGTAT AATGTGATGT GTCTATACAT GTATACATTG
38851 CGGAATAATC AAATCAGAGT GATTAGCAAA TCCCTCAAAT ATTTATTATG
38901 TCCTTGTTGGT GGTGAGAACA TTTAAAATCC TCTTTTAGCT ATTTTGAAT
38951 ATATAATACA TATTATTAAC TGTGGTCATC TTACTGTGCA ATAGAACACC
39001 AGAACTTATT CCTCCTCTGT AAGTTCATAC CCGTTGACTA ATGTCTCCCC
39051 TTTCCCTGTT CACCTCCCCA ACCCCTAGCC TCTGGTAACC CCTATTCTAC
39101 TCTCTACTTC TATGAATTTA ACTCTTTTAG TTCAAGATGT TTTTAAATGT
39151 ACTTTTTTCT TTTAGTTGTT TGTATTCTTT TTTTTTTTTT AATGTAGAAG
39201 AGGCAAATTA AATGCATTAT AAGTTAACAG GAGTTGGTGA TGGTACATTT
39251 ATTTTAACT ACCATGATTG AATTGAATGT GAAACTCATT TTGAATATAA
39301 AACAGCACTA GGTATTCTAT TAGTATTTAT TAGACATTTA TGATCAATTG
39351 ATACTGTCAA TTTGTAATGA TGATCACCAT CTCCAAAAAT AATAATAACA
39401 TCAATTTTTC TTATTACAGT AAAATCCATT ACATGTAAAT TCTAACTACA
39451 GCAAAATTTA GAGCTAGGAT ATTTACCATT CAAGTTATAA TATATCAGAA
39501 ACATCTTATA AAATTATAGC ATTAATTTTT CTTTTCCTTT TCTTTTTTTT
39551 AGGTTAGCTA TTATTTTCCT CTCAAAACCT TATGGAGATC ATTTTTTGCT
39601 GCTTTAGTGG CTGCATTGT TTTGAGGTCC ATCAATCCAT TTGGTAACAG
39651 CCGTCTGGTC CTTTTTTATG TGGAGTATCA TACACCATGG TACCTTTTTG
39701 AACTGTTTCC TTTTATTCTT CTAGGGGTAT TTGGAGGGCT TTGGGGAGCC
39751 TTTTTCATTA GGGCAAATAT TGCTGGGTGT CGTCGACGCA AGTCCACGAA
39801 ATTTGGAAAG TATCCCGTTC TGGAAGTCAT TATTGTTGCA GCCATTACTG
39851 CTGTGATAGC CTTCCCTAAT CCATACACTA GGCTAAACAC CAGTGAAGT
39901 ATCAAAGAGC TTTTACAGA CTGTGGTCCC CTGGAATCCT CTTCTCTTTG
39951 TGACTACAGA AATGACATGA ATGCCAGTAA AATTGTCGAT GACATTCTTG
40001 ATCGTCCAGC AGGCATTGGA GTATATTAG CTATATGGCA GTTATGCCTG
40051 GCACTCATAT TTAATCAT AATGACAGTA TTCACTTTTG GCATCAAGGT
40101 AAGTGCTAAT GTGAGGTGAT ATTTGGGTAA TTTTGGCATG TTCAAACCTT
40151 ATATGTGGA TGAGAGAGGT TGTTGTTTCA TAAATGACTG AAAAAAGTAC
40201 TTATCTTTTG AGTTTAATTT TAAGTAATGA AAAAGATAAT TCCTTAGCAT
40251 ATATTGTTGA CCATGTTATC TGTTGCTATT TAACAAATTA CCCCCAAAA
40301 CTTAGCAGCT TAAGTAACT ACTTATTTTG TTCTTGATAT TGAGTCAACG
40351 ACTTGGGAAG GGCTCAACTG GGCAATTTTT GCTTGTGGTC TTTCATATAG
40401 TTGTTATTAG ACATGGCGAG GGCTAATCAT CTCAAAGCTT CTTTTTTTCG
40451 TTTCCTTTTT AAAAAACTGT TTTTGTGGAT ACACAGTAGC TATATATAGT
40501 TTTGGGGTAT ATGAAGTATT TTGATAGAGG CATGGAGTGC ATAATAATCT
40551 CAGGGTAAAT GGAGTATCCA TCACCTCAAG CATTATCCC TTGTGTTACA
40601 AACAATCCAA TTACTACTT AATTATTTTT AAGTGACAA TTAAATTATT
40651 GAATATAGTT CAAAGACTTC TTCATTCATG ACTAGCACCT AGGCTAAAAA
40701 AATTCAGACA CCTGGGCTCC TGGGATCAAT CACGCATACT GTGTCTCTTG
40751 TGCTCACTCC CGCTGTCTCT CTCTCTTTCT CTCGCTTCCT TTTCTCTCTC
40801 TCTCTGTTG TTTCTAGGGT GGTGGCCTCA GGAATTGGA TTTCTTATAT
40851 TATAGCTCAG GATTCCCAAG AGGGCTGTTT TTAATGTAGC CAAAGAAGTC
40901 TTGCAGCGTG ACTTGTTTTA TTCTATTCAT TGAGGTAGTC ACAGAGGCC

FIGURE 3, page 13 of 27

40951 GACCACATTC AGAGGAGGGA CATACACTTG CTGGGACAAG TGTAAGAGAA
41001 TTCATGATCA TGTTTTTAAAA CCACTTTTAT TAGTTTCCTA TTGCTGCTGT
41051 AATAAATTAC CACAACCTAA TGGCTTAAAA GCCACACAAA TTTAATATCT
41101 TACAGTTCTG CAAATCAAAA GTCTGAAACG GATCTCACTG TGCTAAAATT
41151 AAGGTGTTTCG TAGGGCATTTC TGGAGGCTGT AGGAGAGAGT CTTGTTTTTT
41201 GCCTTTTCTG GCTATTAAAA GCTGCCAGCA TTCCTTGGCT CCTGGCTGTC
41251 TATTTGTCATC TTCAAAGCCA GCAGTAGCTG GTCAAGTCTT TCTCTGTCT
41301 CATCACCTG ACCCAAATC TGCTAAATCT CCCTTCCACA TTTGAAAAAC
41351 CTTTGTGATT ACTTTAGGCC CACGCAGATA AATCAGAAAA TAATCTCCTT
41401 TTTCAAGGTC AGTTGCTTCG AAACTTTCTT TCTGCCACCT TGATTCCTCC
41451 TTGCCATTGCA ACGTAATGTA ATCACAGGTT CTGGGAATTA AGTTATGGAC
41501 ATCTTTGATG AGCCATTATT CTGCCTCATA CCAGTATAGG GTATTAGCTT
41551 GAAAGGACAC TGCAGACTCA GTTAAATTAC TAGATCTATA AATACATGCC
41601 TTTTTCATC AAGAAATTAA GGCAGCTGGG TCTTATGCCC TGGGACATTG
41651 CTTCTTTTGG ATTTATAAAA TAACAAAATT TGTTGATTAA TGGTCTATCA
41701 GTAAATATAA TTTCTTATGT GACTATCAGT GATATATATG GGAAGCACA
41751 TATCAGCTTA TTCTTGTCT TTAATTACT ACCCTGTAC TTCATGTAAT
41801 AGTATTTGCT AGTGATGATG TGCTTTTACA GATGTAAATT AATGTGGAAT
41851 AACAGCTTTG TTTCTACAAA ATTAGAGTGG TTTTAGTTTT TGAAATAAGG
41901 TCTCTTTTCT CTTGTCCTAA GTCTGTAGTC CACTGAGTAT CTAGAGTTAA
41951 ATAATAGAAA AGCCTGGCCA GCGCAGTGG CTCACACCTG TAATCCCAGC
42001 TCTTTGGGAG GCCGAGCGG GCAGATCACA ATGTCAGGAG ATCGAGACCA
42051 TCCTGGCTAA TGCGGTGAAA CCCCCTCTTT ACTAAAAATA CAAAAATTAG
42101 CCAGGCGTGG TGGCAGGTGT CTGTAATCCC GGCTACTCGA GAGGCTGAGG
42151 CAAGAGAATC ACTTAAACCC AGGAGGTGGA GGTTCGAATG AGCCAAGATC
42201 ACACCCACTG CACTCCAGCC CAGGCAACAG GGCAAGACAC TGTCTCAAAA
42251 AATAATAATA AGAAGAAAAT AATAATAGTA ATAGAAAAGC CTAAACATTT
42301 TACCTTTTTT TCTTAGGGAA TCAAGTTAAA AGAGCTGTTA AAGCTCTTTT
42351 TCCTACAATA AGTAAGTGT GGGTAAATCC CAACTTTCTC ACAGTCAGTT
42401 GAACACAAAG AAGCTGGAG CAATTGGCAG GCCTTTGTTA AGTCCCACCT
42451 TTGACTCAGC TCTGGCTGAA GGATCATACC TGGCAAGAGA GTGTAAAACA
42501 CACTTTGATT TTTTCTATTG TTTATCCTTT TAATGATCCT AAGAGACTCA
42551 AGAGTACATG CCATCATTTT GTGTTTGGCT CATTTTCATAT TCAGAGGAGT
42601 TTATTACTCT TTCAGTAGTT TGTTTGTTCG TTTGTTGTT TTTTGAGACA
42651 GGATCTCGCC TTTTTCGCCA GACTAGAGGG CAGTGTGCA GTCTGGCTC
42701 ACTGTAACCT CCACCTCCCA GGTTCAAGCG ATTCTCCTGC CTCAGCCTCC
42751 CAAGTAGCTG GGATTACAGG TGTGGGCCAT CACACCCGGT TAATTTTGT
42801 GTTTTTAGTA GAGATGTGAT TTTGCCATGT TGGCCAGGCT GGTCTGGAAC
42851 TCCTGACCTC AGGTGATCCT TTGGGAGGCC TTGGCTCCC AGAGTGCTAG
42901 GATTATAGGT GTGAGCCACT GAACCTGGCC TCTTTCAGTA GTCTTTAAAT
42951 GATCTTGCTT ATGGTGCTTC TTATCCCTGT TTATTATCCT TATTAAATTT
43001 AATCAATAAA TATTTTCTC TTTTAAATTG ATTCATATAA ATAGACTTAC
43051 CTGAGAGATA TAGGTTCACT TCAGAGCACC ACAATAAAGT GAATATCATA
43101 ATAAAGCAAG TCACATAAAA GTCTTAGTTT CTTAGTGCAT ATAAAGTTT
43151 TGTTTACACT ATGCTGTAGT CTTATGTGTA CAATAGCATT ATGTCTTTTA
43201 AAAAAGTAAT ACTTTAATTT AAAAATACTT GATTGCTAAA AAATGCTAAT
43251 AGTAATCTGA GTCTTCAGTG AATTGTAATC TGTTTGTCTT CTGTAGGGTC
43301 TTGCCTTGAT ATTGGTGGTT GCTAGAGGTA GGACTGGCTG TAGCAATTCT
43351 TAAAATAAGA TAACAGTGAA ATTTGCCGCA TTGATTGACA CTGCCTTTCA
43401 TGAAAGATTT CTCTGTAGCA TGTGATGCTG TTTGATACCA TTTTACCTAC
43451 AGTAGACCTT CTTTTCAAAA TTAGAGTCAT CCTCTCAAAC CCTGCTACTG
43501 CTTTATCAAC TAAGTTTAAAG GAAAATTCAA AATCTTTTGT CCTTTTAAAC
43551 ATGTTTACAA CATCTTTACC AGGACTGGAT TCTACCTCAA GAAACCACTT
43601 TCTTTGTCTA TCCATAAGAA GTAACCTCTT ATACATTCAA GTTTTTTAAA
43651 TGAGATTCTA GCAATTCACT CACATCTTTA GGCTACGCTT ATCATTCTAG
43701 TTCTCTTGCT ATTTCCACCA CTCTGTAGTT ACTTCTTCAA CTGAAGTCTT
43751 GAACCCCTCA GAGTCATTCA TGAGAGTTGG AATCAACTTC TTCCAACTC
43801 CTGTTAATAT TGATATTTG ACCTCCTCCC ATGAAACGTG AATGTCTGG
43851 ATGGCATCTA GAATGGTGAC TACTTTTTGA ACATTTTCAA TTTATTTTGC
43901 CCGGATCAAT CAGAGAAGTT GTTATCAGTG GTGGGTTTCC AAGTTGTCAG
43951 GGGCGAACCA TACAGATCTT CAGCAACCTC AACTCTTGCC TTCTCAGAGG
44001 AAAGAATTCT ACGGAGGGAC ATAAGGCAGA AAAAGAGACT GAGGCAAGTT
44051 TTAGAGCAGG AGTGAAAGTT TATTATTAAA AAGCTTTAGA GTGGGAATGA

FIGURE 3, page 14 of 27

44101 AAAGAAATTA AAATACACTT GAAAGAGGGC CAAGTGGGCA TCTTGGAAGA
44151 CAAGTGCCCC ATTTGACCTT GGACTTAGGG TTTTATATGT TGGCATACTT
44201 CTGGCATCTT GCATCCCTAT TCCATTGATT CTTCTTTTGG GGTGAGTTGC
44251 CCACATGCTC AGTGGCCTGC TAGCACTTGG GAGGGGAGTG TGCACAGTGT
44301 ATTTACTGGA GTTGTATGCA TGCTTACCTG AGGTGTTTGT TGCTTACCAG
44351 CCAAATGTCC CTAGGAGGTC ATATTCATAA ACTCCATGAT TTTGCCTCTA
44401 AATGTGCATG CTTGAGCCCA CTCACCCAAC TCCTGGGATC TTATCGGAAA
44451 GCTGCCGATC GCTAGTTTCA GGTGTTTCTA TCTATTGGAA GATGGCCTTT
44501 CCCTGATGCT GGCTGCAACC AATTATTACT TTAGAGAGAG AGCATGAGAG
44551 CTGTCTCACC ATCATCACCT GATGGTTGCC TGACATTCTT GGTGGGGTTG
44601 GGAGGATGCC TGTCTGCCC TGCTCATGCC TGACTAGCTA CCTGCTGTAA
44651 CAAAAGTACT ATCTATGGTA GCTGTAGCCA TAGGAAATGC ATTTCTTCAG
44701 TAAACTTAA AGTCAAAAT TAGTCTTTAA AACAACATGA ATCTCCTTGT
44751 ACATCTCCAT CAGAGCTCTT GGAAGACCAG GTGCATTATT AGTGATGAGT
44801 AATGTTTTAA AAGGAATCTT TTTGTCTGAG CAGTAGGTCT CAACAGTGGG
44851 CTTAAATAG TTAGTAAACC ATGCTGTAAA CAGATATGCT GTTATCCAGG
44901 CTTTGTATT CCATTTATAG AGCACAGAGA GAGTAGATTG GCATAATTTA
44951 AGGATTACTT AAAAAAAAAG TCTTTGATTA CTCTCAAAA AAAGTCACGT
45001 CTCTCACTTT ATATCAACAG CTAAAAATGG CCAGGTATTG TGGCTCACGC
45051 CTGTAATCTC CATGCTTTGG GAGGCCAAGG CAGAAGGATC ACTTGAGGTC
45101 AGGAGTTAGA GACTAACCTG GGCAACATAG TAAGACCCAT CTCTACAAAA
45151 AAAAAAAAAA AAAAAAGAAA GCCAGGTGTG GTGGTGACAG CCTGTAGTCC
45201 CAGCTACTCA CGAGGCTGAG TCGGCAGGAT CACGCCACGC CAAGAGACGT
45251 GACTTCTGCT TTCAGTTGTA CACTTAGAGA CCATTGTAGG GTTCTTAGTT
45301 GGACTAATTT CAATATCATT GGGTCTCAGG GAATAGGGA GCCTGAGAAG
45351 AGGGAGAGAC AGGGGAACAG CCAGTTAGTG GAGCAGTCAG ACCACATACA
45401 ACACTTATTA AGTTCACCTT CTTCTATGGG CATGGTTCAT GGTGCAGTAA
45451 AACAACGTGA ACAGGAACAT CAAAGATCAT TAATCACAGA GCACTGTAAC
45501 ATATAATAAT AGTGAAAATT TTCAAAGTAT TGAGAGAATT AGCAAAATAT
45551 GATACAGAGA CACAAAGTGA CCACATGCTG TTGGAAGAT AGTGCTGATG
45601 GACTAGCTTG ATGCAAGGAT GTCATAAACC TCAATTGTG AAAACTGCAA
45651 CATGTGTGAA GCACAGTAAC ACAAAGCATA GTAAAACAAG ATATGTCTGT
45701 ATATCAGTCA AAATATTGGG CAACTCTGAT AAGTTTGTCC ACTTAACATT
45751 GTACCACTTA AGATGAATAG CATCTACCAT TTCCGTCATT TGTAATATA
45801 TAGGAGGACA TAATCACATA ATCTTGAAGT AAAAGACAGT GCTTAAACT
45851 GAATCAGTTA AGTTTTATGA AAAACTCTC ATATTGTAAT TTTAAAAATA
45901 TATATTTTTT AATTTCAATA GCTTTTGGG TACAAGTGGT TTTGGTTACG
45951 TGGATGAATT CTATAATGGT GAAGTCTAAG ATTTTACTGC AACTGTCACC
46001 CAAGTAGTAT ATATTGTATC CAGCATATTG TCCTTTTTTT TTTCTTTTTT
46051 TTTTTTCATT TCACCATGGA CTAATGAAAA TTTTGTAGG GACTGACATT
46101 AGGGCACCTT TGAGCTACCT TGAGCTAAAG GAAATAACCC TTGAATTTTT
46151 TTCTGTTTGG CCTAGAGAAT GTGGTTTGT TTTGTAAGT ATTATGTTG
46201 TTGTAAAGGT ACAAGATTTT GCTTTAGTTT TATTGTACT AGGATTTTGC
46251 TATATTAATA CAATGTGAAA AGAATCAAAA GTGTAGAAA TAAATGCATA
46301 GAATGTAAGT TTCAGGCATG TGAGTAGAGG ATCTCTGCTC CATAAAGAGT
46351 TCTGTTGTG TTATAGGTTT CATCAGGCTT GTTCATCCCC AGCATGGCCA
46401 TTGGAGCGAT CGCAGGAAGG ATTGTGGGGA TTGCGGTGGA GCAGCTTGCC
46451 TACTATCACC ACGACTGGTT TATCTTTAAG GAGTGGTGTG AGGTCGGGGC
46501 TGATTGCATT ACACCTGGCC TTTATGCCAT GGTGGTGCT GCTGCATGCT
46551 TAGGTAATAT GGCTGTGTCT GCCTGTGTG GGATGTTTG AAGTCTGAGA
46601 GAGCCAAGAG AAAGTGGGAC ACATTCTTGC TTAATTGGTG GCGGATTGG
46651 TTGAGTAAAG GAGGGTGCCA GGAGGAGATG TTTTAACAGA TAAGAAACAG
46701 TAGTACTATT AGGGTATTAT ACAGTACCGG TTTTCTGTCT TACAACATTT
46751 GTTAATACAA GAATTTAATG GCATTAGCAT ATTGTAATAT AACTTAATAC
46801 ACTATGGCAG AAGCCATCTA AGTACAACAT AAGCTTAATT TGAATCCTGA
46851 CCAAAGATGT CTTTGATTCT TTCATCGTTA AGGATCTTGG CTTACCTATA
46901 ACAACTATAG CATAATACCT AAGATTAGCA TTGCAACAGA GTTTCAGAGT
46951 AGGTTTACTT TGGTTCTGAA ATGATTATT GTTAGCCTTA GTAAAAGATG
47001 TATTTACCCA TGCTCCATCA TCTAAGGTAT ATTTGTAACA AAATGAGAAA
47051 AGGTAACCTT ATTTTAATGA GAAGAAAAGC AAAATACCTA CATTAGTAC
47101 TTGAGTCTAT TTAATGCTG TTAGGGCAGG AAAAAATGGT TATTGCTTTT
47151 CATATTTAAA ATATCAGCTA CACTCTGGTG ATAATATTAA TGGTTGCCAT
47201 TTTGACCACT TTTGTTTGTG GAATAAAAAT TATGTGATTA TTGATCTTTA

FIGURE 3, page 15 of 27

47251 AAAATGTAAT ATCAATTAAA AGGAAAGGAC AGACTCATTT TCACCAAAGT
47301 AGCAAGTATT TATTTAAATGT CCACCTTCTT TTTAGCATTG TGCTAGATAC
47351 AGTGCATAAT AAAAAAGAA CATGGACCCA ATCTCGACTC TAATCAAGTT
47401 GAGGAGACAA GATGAACACT GAGAATACAA TAGTGAGGAA TACTAACAAA
47451 TATATACAAG GTTAAAAGAG TCTAAGTATG GTAGGAATAT AGGGGAAGAA
47501 AGAGCTGAAG TACTTCAGGA AGAGTAGAAC ATGAGGCTTT ATTTAAAAGA
47551 TTAGCAGAAT TTAAGGAAAA GGTGACTTTG TTGAAGATTA TAATGTGAAG
47601 ACAAAGGAAC GAGGATGGGA ATAAATTTTG TATTCATGAG GCTTTGAAGA
47651 AATTGACTCT AGAGAGTATA TTTTGGGTAC TTTTGGGAAA TGAAGTTGGA
47701 TTAGTGAGAA GGAACAGATT ATGAAAAGAC AAGAAACCTG ATTAATGTCA
47751 GGATGATTTT ATATTTGAAG TTGGTCAGAT TTATGGCAGT CCTGGCTTTG
47801 CCATTTTTAG TTTGATGACT TTGAGAAAGT TCCTTCTTGA AGTTTAAATT
47851 TTCTGTATAT AAAAAGTAAT AACACCTGGT GATCTGCTAG GTTTGTTTTG
47901 AGGATTATAT GAGATAAAAT GCATGCAAAA CTGTTATAAT AGTGCCTGGT
47951 AAAATAAGTG CCTAGTTTTA AAAACAAGTC TTTGTAAACT GCTTAGGACA
48001 TGCCTGGTAT AGGGTAGGTA TGTAATACAT AGTAGGTAGG ATCTGTCTCC
48051 TTGCTATTTT TAGGTAAAAA AACAAAAGGA AGAGCTTCAG CTTAATACAG
48101 TATGAACTGA CGAGCCTGG TAGGTTTTTG AGCAAAAGAG CAACACAGTA
48151 AAAGTAGTAC TTAGGAAAAGA TTAACAAGGG AACATGGCTT ATACAGTGGT
48201 AATGGGGCCT GGAGTCAAGG AGGTAAGATA AAATGGTATT ATAATTAAGG
48251 AATAGCCAGG CACGATGGCA CATGCATGTA ATGCCAGCTA CTGGAGAGGC
48301 TGAGGTGGGA GGATCATGGG AGTCCAGGAG TTTGAGACCA GCCTGGGCAA
48351 CTGAGTGAGA CCCCCAATCC TAAAAAATAC AAAGTAAAAA AGGAATAAAG
48401 TCATGAGGGC TTGGACTGGA TTGATAACAG TGAGAATACC GAGAAAGGGA
48451 CCATAGGCAG TGTGAACGCA GCTCACTGCA GCCTCAAACC CCAGCCCAAA
48501 CGAGCCTCCC ACCTCAGCCT CCCAAGTAGC TGGGACCACA GACATACACC
48551 ACCATGCATG ACTACTTTTT TTAGTTTTTA CTTTTGTAGA GACAGGGTCT
48601 CACTGTATTG CCCAGGCTGG TCTCAAATC CTTGACTTAA GTGATCTTCC
48651 TGCCTTGCC TCCCAAAGTG ATTACAGGCA TGAGCCACAG TGCTTGCCCC
48701 AAATAGTTTT CTGTGAGTGA ATATTACTTG CATCGTTAAT GTAAATCAAA
48751 GGCATCAAAG TGTTTTACTC TTTTGGAAAA AAATTAGAG GAGAAATTTA
48801 TTATATTAAT ATTCTACCCA TATATGAGTT TAATTTGTAA ATTGTAGCAA
48851 AGCATGATGT GCTTTACTAA ATTCTTTTAT AATTAGAATA AGCTTTTATA
48901 AGGGTGAAAT TATGTCTTTG CTACAGCACT AAACCAAAT GGCAAATTTG
48951 TTTTAGTCGG TAAGCTTTGC TTTTAAAAA TATGAAATAA ACAGGTTTTT
49001 AAAATGTTAT TTTAATAGTC TTCTCTGTTA TAAACAAAGA AAATTGGTGT
49051 TTCTCTAGAG CTTATTAATA GTAGTGATTA TTGTCCTAAA AGAGGAGTAG
49101 CAGTTTTAGA TGCTAATGCT TTTCCCTGAC TGAGTTCTAT TTGCCATTTA
49151 GTTTTAACTG CCTAGTGCAA AAATTCTAAT AAAATGTAAT GATGAGGATC
49201 CTGTCTTCC TGACCAAGTG GTGCTTACTT TTTTCAGGTG GTGTGACAAG
49251 AATGACTGTC TCCCTGGTGG TTATGTGTTT TGAGCTTACT GGAGGCTTGG
49301 AATATATTGT TCCCCTTATG GCTGCAGTCA TGACCAAGTAA ATGGGTTGGA
49351 GATGCCTTTG GCAGGGAAGG CATTATGAA GCACACATCC GATTAAATGG
49401 ATACCCCTTC TTGGATGCAA AAGAAGAATT CACTCATACC ACCCTGGCTG
49451 CTGACGTTAT GAGACCTCTA AGGAATGATC CTCCCTTAGC TGTCTGACA
49501 CAGGACAATA TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACCAG
49551 CTACAATGGA TTTCTGTCA TAATGTCAA AGAATCTCAG AGATTAGTGG
49601 GATTTGCCCT CAGAAGAGAC CTGACAATTG CAATAGGTAC CCTTCAAAA
49651 ATATATATAT GTATATATGA GATGGATTTC TGGAAGAAAG GAAAGCAATA
49701 AGCAGTAACA TTTAATGGGT CGGATTTGTG GGGCAAGGG ACATTATTTT
49751 ATGTCCCTTA ACATCTTCTG TTCTTTAAGA AAGGAAGGTA TGCTTCAGTG
49801 GATGATTTTC TGCTATATAT CACAAAATCT GTATTTCAGG TTTGTCTTTT
49851 GATCCGGCAT GTACCAGAAA TTGGAGTCAG ATTATTTTCC CACTCAGATA
49901 AGCCTAGATA AGTTGATCTT GGTATTTCAA AACAGCATGT AATATAAGAC
49951 CTTAGCTAAA TGCAATTCAGT CAAATACATT CTTGTATTTA ATAAAGTTGG
50001 CTTATTGGAA TACAAGTTAT TGAAAAATCTC ATCTTCATCA GTCTCTTTCA
50051 TATTAGAATA ACTAGTTTT GCTTTATCAG TCTTTGGGGT TAGAATTATA
50101 ATATTAAATT ATAATATCTG ATTTAAAGTG ACAATCACTG AGATTTTTAT
50151 TTCTGATCAA ATGCCAGGTT GAAAAAGTAT AACGTATCAG TCCTGTTGTG
50201 TTTTATGCAG ACTTTCCTGA AAATACTGTT TAAAGGTATT AGCCATAGTG
50251 TATTTCTTGG AGATAAATTA AACTTTCTAT AGTTCTGTTT CTCTAAATTT
50301 TGTTTTTCTC TTTACCTTAT AGTCCCGCAG TATTGATGAG GAGACCATTA
50351 AGACTTAATA TTTTTTTGAC ACAATCTTAT ATCTCTTCT CCAACCCCTA

FIGURE 3, page 16 of 27

50401 AAAAGTGA CT GAGGATAGGT ACATCAAGCC ATTGCTTTGT TACTCCCCAG
50451 GTTTTAGTGC CAGACCCTGA ATGGAAGTGT CAAGCCTTTG GCCTGTCTGA
50501 AAGGTCATTC CTGTGAGCAT ATCATCTCCC TTCCAGCTTA CCTCTGTGGC
50551 CATTGCAAAA GGATTTAAAA ATAATTTTGT TGCCATTGTA ATGGCACAAG
50601 ACCAGACAGT GTATGTGGGG GAGTGTCTCT CAAATCAAAAC TGGAAACTCT
50651 TTAATTTGTA AGAACCATTA AGCAGAGAGA GAAAAAAGAA AGGAAAAGAA
50701 AAAAGATCCT ACAGAGAACA CCCTGTTCAG TTTGGGAACA GGCTACAGCT
50751 TTGGATTTT CAAGGCCTAG CATTCCCATC ATTCTAAATT TTACTTAGCT
50801 AATACAATAG TAGTTGCCAG AGCTGATGAC ATAGTATTTT GTCATGCTTG
50851 GCTCCGTTCA AGCATTTTAG TTTTCTAGCC ATTACCATGG CTAGACCCAG
50901 TCAAAAGAA TTTTCATTGT TAAGATTCCC ATTATCCTAG TTTTACTAG
50951 TAGCCAGCCA AAGAAAAGAA AAAGGAGGTC AGAATTTCTG TATTTACATA
51001 GAAATTTAAG GGGAAAAGGC CAGGCATGTT TTTAAAGTGT GGAAATTAAG
51051 AACTATTCAT TATCCCACTG ATTGTGTGGA TGTGTTTTTT AAAGTTTGT
51101 TACTGTCTTG AGAGAGAGAA TATTGAGATA GGACATAATG TTGGTTTAAG
51151 GGAATGAGGG TACTTTCTGT AGGTGAGGTG CCAAGCCATG TCATCAGAAA
51201 TGTTAGTCAC ATGACTTTCT AAGCACACCT TAAATGTTTT ACCGTGTATG
51251 TTTTGTGAAA GTTTTAAATT TTTAACTGGG AAAAAACAGC CTGTATATTA
51301 AGTTTTATAT ATATATATAA ATTTAAATTT ACATATATAT GTTTATATAT
51351 GTAACTTTTA TATGGGAGAG ATATATATTT CTATATCCTC TATAAAAAAA
51401 CATATCTATA TATGAAAATT ATGTACGTAA ATGTTAATTT ATAATTAATT
51451 ATATAAATAT TAACATAATT ACATTATATA TATAGAAAAC CTAGTGTA
51501 GATCTGTATA TAAATTAATA ATGTATGTGT TATATATAGT TACATCATAT
51551 AATACATATA ATTGATATAT ATAATGATAA ATACTTTATT GAAGGATGAA
51601 AAAATTTCCA TGCTGTCTCA TAAAATAAGA TGGTTGACAT ATGCTAAACT
51651 AGATAGATTC TCCTGTTTCA TACTAAAGCA GAATGTTGTA AAATATTAAA
51701 TCCAAATGAG ATGTCTCAGA TTAAGGCCAT TTCAACAGGA ATGCTGAGAC
51751 TTTAAAAAAA AAAAAAGTCT GAGGCTGGGC GTGGTGGCTC ATGCCTGTAA
51801 TCCCAGCACT TTGGGAAGCT GAAGCAGGTG GATCACTTGA GGCCAGGAGT
51851 TTGAGACCAG CCTGGCCAAT GTGGTGAAAT CCCGCTCTA CTAAAAATACA
51901 AAAAAAATAC ATGGGTGTGG TGACGCATGC CTATAATTCC AGCTACTTGG
51951 GAGGCTGAGG CAGGAGAATC ACTTGAACCT GGGAGGTGGA GATTGCAGTA
52001 AGCCCCACCA CTGCACTCCA GCCTGGGCGA AGAGCAAAAC CCTGTCTCAA
52051 AAAAAAATAA AGCCTGAATT ATATCAGCAA ATGAAAACCT TAATGTTGTT
52101 CTCTGTTTCA GAGGCCCTTG AATGAATAGC ACTAAAAATA TTTTAAAAAA
52151 AATGAAGAAA ATGAAAATTG TAATGTTTCT TATTTAAAAA GCCCTTGAAT
52201 GAGTAGCATC AAAAAATATT TTAATGGGGA GGCCAGGGTG GGAGGTTTGT
52251 TTGGCACCAG GAGTCAAGA CCAGCTTGGG TAACATAGCA AGACCTTTGT
52301 CTCTACCAAA AAAAAAAT TGGGTGTGGT GGTGCCACCT GTATTCTTAG
52351 CTACTGGGAA CACTGATGCA GGAGGATCCC TGGGACTCTA GAGTCCAGAG
52401 TGAGACCCTG TCTCTAAAAC AAACAAACAA ACAAACCTG TATTTATGTA
52451 AAAGTAATAC TTGTTTTTTA AATTTTATTT ATTTTAAATT GATAAAAAAT
52501 GTATGTATGT TTATGTGATG TATATATTGT GGAATGGTTA AATCAGGCTA
52551 ATTAACCTAG ATTTTGTGT TGTGTGGGGA GAATATCTAA AATCCCTCTC
52601 CTTAGCAGTT TCCAAATGAA ATGAAAGAAT AAAAGTGATT TATTTTGTG
52651 AGACAGCATC TCACCTGTT TCTCAGGCTG GAATGCAGTG GCACGATCTT
52701 GGCTTACTTG ATCCTCGACT TCCCTGGCAT CCGGTGATCC TCCCACTTCA
52751 CTCTCCTAAT TAGCTAGGAC TACAGGCATG CGCCACCATG ACTGGCTAAT
52801 TTTTGTATTT CTGTATAGG CAAGGTTTTG CCATGTTGCC CAGGCTGGTT
52851 TCAAGCTCCT GGGCTCAAAC GATCCACCTG CCTCAGCCTC CTGAAGTGCT
52901 GGGATTACAA GTGTGAGCCA CCACACCTGG CGAAAAGTGT TATTTTAAAA
52951 AATGACAAAT TTAAGTCAAA GAGATTGAAT GTTCACTTCT GGTACTTTGT
53001 ATATAAGAGA AACATTCCAT TAAATAATTT TTTAAACATT TCTAAAATTA
53051 CATATTTTGT CATTAATATG TTAACAATC AGTATAATTT CATTGATACA
53101 GTGTTTGTGA TTTTGTCTGG GTTTAAGATT GATAATTGGG GTTAGTTTTA
53151 ATTCAGAAAT TTATCTATT TAATGTCACA CTTCATGTCT TTTTATTTTG
53201 TATATCTATT TTTAGCTATA GTTATTACTG TTTTAGAGAT TTTTAGAGAT
53251 GAGGTCTTCT ATGTTGCCCA GGGTAGACTT GAACTCCTGG GCTTCAGCAA
53301 TCCCCTCCTC AACCTCCGGA GCACATGAGA TTAGAGACGT GTGCCACTGT
53351 ATCTGGCCTG CTGTAGTTAT TTTTAATTCT TTTGTCTTTC AACTTTTATA
53401 CTAGAGTTAG AAATGATTTA CAAACCCTAT TGCAGTTTTA GAGCGTTATG
53451 AATTTGACTA TATATTCTTT ATAACAACCT AACTTCAGTT GCTTACAAAA
53501 ACTACAGAGT TTTACTCCCC CGTCCACATT TTATACTATT GATGTCACAC

FIGURE 3, page 17 of 27

53551 TTTACATCTT TTTATTTTGT GAATCCATTA ATGATACTTC TGGTAGTTTT
53601 TACACTCCAC TATTCAGTTG TCAGACACCA TTCAGTTGTT AGATTGTTAT
53651 GAGCTAAAAG CAACCTAATG GGTATTTTTC AAAAATCATT TATGTCAATT
53701 GCTAATGGAC TTCTTTTCTA TGCCATGATC ATGCTTTTTT TATTTTGTAG
53751 ACGGAGTTTC ACTCTTGTG CCTGGGCTGG AGTGCAATGG CGCGGCCTCA
53801 GCTCACTGCA ACCTCCGCCT CCTGGGTTCA AGCGATTCTC CTGCCTCAGC
53851 TGGGATTACA GGCATGTGCC ACCGTGCCGG CTAATTTTGT ATTTTGTAGTA
53901 GAGACAGGGT TTCACCATGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC
53951 AGTTGATCTG CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGACGTGA
54001 GCCACTGCGC CTGGCCTGAT CATGCTTTTA AGGTGGTTGA GTAAGTACTA
54051 GTTGTGGGG CTATAGTAGT TGCCCTCCTA CTCAAATGTG TTAGAACATA
54101 GTTAAGAAGG CTGTAGTGT CAAAAGGAGT AAAAAGCAGT GCAGTGTTTG
54151 CAGTAATATC TGCTTCTCAA TTTAGGACTG ATGCTTATTA TGGCTTAAAT
54201 GTTTTTGTAG TAAATTTTGT ATTCAAAAAA TATATTTTTT TTTCTTTTGT
54251 CGACAGAGTC TTGCTTTGTC ACCCAGGCTG GAGTGTGGTG GTATGATCAT
54301 GGCTGACTGC AGCCCTGACC TTCCGGGCTC AAGTGATCTT TCCACCTCAG
54351 CCTCCCAATT ACTTGGGACC ACCAGCATGC TTGGCCGATT TTTTTTTTTT
54401 TTTTTTTTTT GTAGAAGCAA GGTTCCTTA TGTGCCAAG GCTGGTCTTG
54451 AACTTTAGGG CTCATGTGAT ACTCCTGCCT CGGCCTCCCA AAGTGTTAGG
54501 ATTACAAGCC TGAGCCACCA TGGCCGGCCA AAATATTTTC ACTATAACAA
54551 ATATCATATC TGTATATACT CAGTTTAAAT ACTAACTCAA AGTAGAAACA
54601 TAAAGCTGAA TGACTATTTT ATTTTCAGAT TCTCTCCATT GAGTTTCCTT
54651 CTCCGCTCTG TGTGATCTCT GAACTTTTCT CCATCTTTGC CACTTCTTGT
54701 CTAGCATTTT GTTTTTATCA GCAGTTTCAT TCAGATTTTT TTTTGTAGTT
54751 TTTCAACGGT GGAGTGGGAG TAGGCAGCAG GACAGAAGAA CTTGAAGCAG
54801 AGCACACTGG AGAGGAGAAA TTAACAAAGC CTTTATGAAT AAAACAACCC
54851 CCCAATATCA GTCTGTGTGC ATTATGAGCA TAATGTACTT TTCATCTCAT
54901 CTGTAATGTT CATGACTTTT CTAGAAAATT ATACTTTAAC ATGAGAAAAG
54951 AAAAAGAACC AGCTAATTCA TAGGGATGGA GGACACAGCA TAGTCAAAGC
55001 AAGAATGAAA CTCTCTTTAG TGCCACCTCC AGTGCAGAA AAGTAACATT
55051 CAGCAGAGGC AGGTTTCATT TGATAATGGA TTCTTATAAT AAAGTGCCT
55101 CAGAATTTGT GCAGGTTTTA AAATCCCGTA TTCCAAACCC ACTTCCCTAG
55151 CCCCCAAGTT AGAAAACAGC TTCAGTAAAG AAAATTGTAC GATGATATAA
55201 CTTTACCAAA AAATAATTTC TTTCCATGAA GATGATATAT TATTGTTGAC
55251 TTCTAATTCA ATCAAATATA ACAAATGCT AAATGGCTTT TCAGTTGACT
55301 CCTTCTTTGG TTAAGGAGAA GATAGGAAAA AATGAAGGGA TCAGAAGTCA
55351 TAGGATACAT TAATTTTTTT TATCTCTGAA TAAACAGGTT GCCTACTTAA
55401 AAATCTATCA GTTTTAAAGT GTTGGTCTCT TCTCTCTCTT TTCAGAAAGT
55451 GCCAGGAAAA AACAAGAAGG TATCGTTGGC AGTTCTCGGG TGTGTTTTGC
55501 ACAGCACACC CCATCTCTTC CAGCAGAAAG TCCTCGGCCA TTGAAGCTTC
55551 GAAGCATTCT TGACATGAGC CCTTTTACAG TGACAGACCA CACCCCAATG
55601 GAGATCGTGG TGGATATTTT CCGAAAGCTG GGAAGTGGC AGTGCCCTGT
55651 AACTCACAAAT GGGTAAGTCT GGTACCACAG GAATCAGTTC ACTTGCTAGA
55701 ATATAGGATC CTTTTAGTGT GAATCTATAT AGTTATTAGG GGAGCATGTG
55751 AGTCAGCTCC CAGGTGGGAA AGTCTGTCT ATGGTATAGT CACAAATATA
55801 GGATCAGTCA ATCAAATTTT ACATTTACTA AGGAATAAGA AAGATGTCAT
55851 CTGCCTGCTC TTTGCCAAAC AGTGACATTT GTAAATAATA CCTCAAAGTT
55901 GGAAAAGAGG TGCTGAAAGA TCTCCAGCAT GAAAGCATGT TGAGCTTAGA
55951 GTGCTTCTTT TCCTAGGGAA GAGTGGACCT AACCTGCATG GAGCACTGCA
56001 AAAACCTGTT TTATTTTGT AAATGTTTCA TTTTGTAGT ATAAATTTCT
56051 AGTACAATAA TAAGTTTCTA GATATTTTGC TATTTACTCT TTCAGCCAAT
56101 ATTTGATTTA TCATGTAATG AAGGAAAGAA TATATACTTA AATGAAATTT
56151 GTAAATGAGC TAAAAATCTC CTTTAAACAA TGCTTTGTTT CCTTTTGTCT
56201 ACCTTTCTCT ATACACAAAT CTTTTATATT TATATACTG CTAAGGACAA
56251 ATAAATACTC ATGTATTTAA AATGTATACA TTGATAATTT ATTTTCCAC
56301 CTTTACACA TGAAGTGGCA GTGTTTCTCC ATTGACAGGA ATATAGGAAA
56351 GAAACAGATG TCACGGGGT TGTGGAGACC TTAATGCACA GAATTGATTT
56401 AGCAAATACA CTACTTCGTC ACCACTGCTC TCTTTTCTG GACCTGGGAT
56451 CTGTTTCTCC ACATTCTTT CTTTAGGACC CTTTATTTCC ACTATATATT
56501 CTTTCTTGTG GAACTTAAGA ATGTTGTTTT ATCCGAAGGC AAATACCAAA
56551 AAACAGAGGG TATTCTTGGA TTATGCATAA ACTGGATGGC TAATCCTGAA
56601 CAGCGTAAAG CTGGTTGAAA TTCTAAACAG AGAATCATAG CAGTTTTTTG
56651 TTGTTTTTTT TTTTAAACAT GTTGTAGAAA ACACATTGGT GACAGAATAC

FIGURE 3, page 18 of 27

56701 ATGACTCCTG TCCAGAGAAA GGAGAGAAAA AGAACAGAAA GGAAGGAAAT
56751 TTGTTTATTG AACACCTTCA TATTTTCTCA TTTAACTTTG CAGGACCTCT
56801 GCAAAGTAGG TAGTTATATC CCTACTTTAC AGATGTAGTA ATTAAAGCTC
56851 AGGAAGCTTT AATAATTTGC CCAAAGTCAT GTGGTGAACA AGTCATGGTT
56901 CAAGGAATCA GACTGTCTTT CCTACTTTAA AACCCAGCCT CTTGCTACTA
56951 TTTTGCAC TG TAAGTGACTG ATAGAAATCC TCTTTCTTTG TGATTTCTTA
57001 AACTACTAAA ACATTTTCTT GGCCAATATA TTAGATTGAG TTAAGAATAG
57051 AAATATGAAA CTAGAGAATT AGATCTATGT TTAGTGTTTT TCACTGCGCT
57101 AATTAAAAATA ACTCTTTAGG AATATGAAGT AAATCATTAAG AGAGATAAAG
57151 CCCTTAAAGG CAGGGAGTTT AGAATTATTA AATTCTAATA ATTTAGATAC
57201 TGATTGGAGA AGAGATGTAT TCATAAGTTA TTATTGTTAC TATTTGTCTT
57251 TGTGTAATAT TGTTTGATTA AATGATGGCA CCGACTTCTA TAAGTTTAAA
57301 AACTCAGTAC TAGTTAAATG GGGCAACTTT TCATAAAGCT TTGCTAGTCC
57351 TTGAGCCCTT TTATTTGTGA AATGGCTCAA CTGGAACCTA AGCTGAGTTG
57401 TTACAAACTA TTATTTGCTT CAAGTTGTTT TCTGTCCCTG GCATGGCTTT
57451 TTCTTTTGTG TACTGACAAA TATAAATGTT ATTCTGTTGA GTTATGGTTA
57501 ACTATGAACA CAGAACTGTT AGGGATTAAT TTTCATATTT CAGTTTGTG
57551 ATTAATTCCC AGGTATTTGG CAGCATAGAT ATTAGAAAGG AAAATATTTA
57601 AAAGAAAGTG TAAAAATAAC GAAGTGATA GAGCGAGGGG TGGATAGCTA
57651 ATTAATAATTT TGTCTGGTCC TGCCTGTTCA TATGAAAAAA GGGGTGGAC
57701 TTTCTTCTAA GGGAATATAT TAAATTGCTT TCATCATATT TTCCTTATTT
57751 CTGTCTGTCA AGGAAAATAA ATTGATACAT ATATGGGGAG AAAAGAGATC
57801 ATTTAGGGAA GTGGCTCATG GGACTTTTTG TTTTGTGTTGA AGTGATTAG
57851 GAAGTCGGGT GTTTTTTTT TCACTTAAAT TATTTAAAAA CCAGAAAAGA
57901 AATGATATCT TCTGGTTTTT AAAGGAGACC ATGAAGTTCT GCATAGCTAT
57951 CATTGATGTG TAGTTCATAC TGCATTTTTA GAAGTGGAAA ATAGTTATTT
58001 GGAGGAAGAT AACAACTCTG GAACCTTAGG TGCAAGGAGA AAAAGAATAG
58051 ATGAAAGGGA AAGATGTTTG TAAATTATAA AAATTTC AAT TAGCTATTGG
58101 TTTTCTGCAC TTTATATTTT AACTGCAGAA TTTTTCAAAA TCAGTTAATC
58151 TTGGTGGAAAT TAGCAGGATG TTAATAGGAG TGAATCAGAA AAAACATTT
58201 TGTGTAATCT TAAGTTTGGG AAGTATTGGA TTAAATACAA TTGAGGTTTC
58251 TTTACTATGG AACTCCTCAG AACTTATAAT ATGTTGATAT TCTTTGATTC
58301 CCAGATGAGG GGATGGGTAA TAGGATACAT GGTTTTCCAG ACTTGTGTTGA
58351 AAATGCAACT ATTTTGGGT TGCAGGGAAG GATATAGTAG AACTCATGGG
58401 AACTGGTGT TCTTGGACA TGCTTTGGAA ATGCTGGGT ATGCCTGTT
58451 AACTCTTACA TCATTAGTTT TTAGCCCAAA AGGAAACAGC AAATAATGTT
58501 TTATATGAGC CACATTTTGT GTTGATTTTC CTCCACTCT GTAAATTTAC
58551 TAAAGCAGCA CTCTGACTTT ATTATGCTCA AATCGCTCT CTCCATTAAT
58601 GTGTGTTTCT CCATCTTTTA GGGTTTTTAC TTTATAAATA CAGAGATTAC
58651 TGTGTAATAAT TCTAAATTTG CCACTGGGTC GTTATACATT TGTAACCTTC
58701 CTCACAGTAT ATTTTGTGAT TTGGCAGAGT TTACCAATAT AGATGATACT
58751 AACTGAAATT AATCATTTCT TATAATTGGA TAGAAAAGCA TGAGTAAGAA
58801 TTCAATTGGT ATTATATTTA ATTAATTGCC AAGATTTTCA CATTTCTCTGA
58851 CTACAACAAT AAAATCAAAT GAATTGATGG CTTAAAAAAA AGAAATCTCA
58901 AATGTTTATG CAATGAAGAA CATCTATTGA ATGAGTGAAT GTTCATTATA
58951 TATAGTGCAT TTTCTGAGCT TTTTGGAGG GGGAGTTGC TCCCATGCTC
59001 TGAGAACTTT TAAGGATCGA TACATTATTT TTAACATAAT AATGAGAAAA
59051 CATGAGCAGA GAACCCATTT CTGTCAATCC CATTCTCTAT CCTCTGCTC
59101 CCCCACCTCC CACCCAGCC ATCAAGCTAA GTAACATTTT TACACCTGGA
59151 CGTAGCTATA GGAACAGGCT ACTTTGAAGT CTCCTAGTGA CATCCTTCAA
59201 GTCTGAATGT TCAAAGGCAG TTTAACAGGG AGGTTGACTT AATGAGATCA
59251 TCAAGGAAAT GTCCAGTCAT CCTGAAGGGT ATTTTGGATG GGCTTCCAGA
59301 ATTTAAAGAT TAAAGTTTTT TTAAGGTTTT TTTATTTTCA CTGTTTATAT
59351 TGCCACATTA ATTTCCATTA TAAAACCAAGT AACCATAGTT TTGTTTTAAT
59401 TAGCAATCTA ATTATTTTCA TGTATCTCTA TTATGAGAAT TTATGTCCAT
59451 CACTTTGCTT GATGTGATAA CAGTGACATG CTAAATGAGA AACAATTGTT
59501 ATTTAGAAAA AAATGCACAA AGTGAAGTC CTTTAAATCC CTAATCATAA
59551 ATACATTTTA TTAGCTTACT TTAAGAAGTG GCAGTCACAG CTCCTGAACA
59601 TTAGGGAGTG TTTCTTTTGG TCAGCATTAT TTATTTAGTG CACATTGCCT
59651 TTAATTTTAA TTTGAAATTA TAGTAAAAATC CACGGGAGTT TTTAAGTCTC
59701 CTCACAGCCT TTTGTACCT TTTACCAAG GTAGATCCAG ATGATAACTG
59751 CTGTGTTGTG ACATCATAGA AATTAGAAAA ATATTTTCTT CTGAGGAAAG
59801 AACATTGTAA ATGAACTCT ACATATCAGA GGTCTATAGC TATGTATCAA

FIGURE 3, page 19 of 27

59851 TATTAAGTTT CTTTTGTACT TTGCTTTGTA GTCATCTTCA TTCCAAACTT
59901 TCATAATTAT TATTTTTTACT TTAAAAAGAA AAATAACCCA CCAATATTGA
59951 AGATTAGTAT TGTGTCACCT TTGAAAGTCA GTAGAATTTA TGCAAAAGGA
60001 ACCTGGAAC TTAATCATTT TTGTTTTTAT TTTCTAAAGT TCATGAGACT
60051 CATTCTTATG GTTCATGTTT TTATTTTTTC TCTCATCTCT TATCATATTG
60101 ATTGGAACCT CTTTTAATTT AATTTCTCAC ACAGTTATTA GCATAATAAT
60151 CTGTTTCAGG ATTGTCTTGG GGATCATCAC AAAGAAGAAC ATATTAGAGC
60201 ATCTCGAGCA ACTAAAGCAG CACGTCGAAC CCTTGGTGAT TAGATATATC
60251 AGATCTCCTC ATTAGACACC TTAGAAGTCA GGAAGCATGA AACTTGTGAA
60301 CTGTTGAGTT CTGCTTTTCC CAGATATCTG CTGAACAAAA ATATCCTACT
60351 ATGCTGCCAA TTACATTTGT ATCTGATAAA ATGTGCTGT AAGATAAATT
60401 TAGATATGTG TAAATCCCA TTTATAGAAA GTAAGCAAAA GTTAACATCT
60451 CTCATCAAA CATTCAATAC AATTTCTAGAA CTGTAAACAG TTTGGTAGTG
60501 GAATAAGTGA ATATTATTGG ACATTCTTAA AGTGAATATG GCAAATCTGT
60551 CTACCTCAGT GGATACACCG GTCTCAGAAG ACACCTGACT GGTAAAAAAT
60601 GTCTGACCCA TCCCCGCAAG CCCTTTTTTT TTTTTTTAAA TGTTTCCCGA
60651 TCTTTGGTAT GTCTTATGGT AAATCTAAGC TCCTAAAGGA TTTTAAAGGA
60701 GCTTAGCAAT TAGAAGTCTG TACAGTTAAA TGGATTTTTT AATGGGCACA
60751 CTAAGTAGAG TGTAATGTGT ATATTATTGG TGATCATAGC ATTAGTTCTT
60801 TTTCTGCTAT ACCCTGCATA TCTTCAAAGT CACAGTGTGT GTCCTGCCAT
60851 CTCATTAGTG AATTGTACCT AGATTATGTG TGTGCCCTTT TTGTATGATG
60901 TTTCTGGAAC GCTATAAGCA GCTTTTAGAG TCAAATGCAT TCATTTTAAC
60951 TGGCTTTATG TCCTAGTGGT TTCATGACTA CAAATTTGAA TTATCTTACT
61001 GCATAACATA AAAAAATGCT GGCTTTAGCA ATTAATGCCC GAAATATTAT
61051 TGCCCTGCAA TTGTCATACC TGTATGAAAC CTGTCCCAGT TTGCTTAAAGT
61101 GCACAACTGA TTATGTATTC CTGTGTGTAT GCTAATATTT CACAAGTGTT
61151 TCATGCATCC TTTTTTAAAA AACTACTAAC CAGAATATTA TCGTAGCTAC
61201 TCATTCATTC TGCTTTCTGC TTCACCTATA ATAATCTTTT AGGACTGCCT
61251 TCTGATTTTT CACCTATCTT TTAATGTAAG CATTAACAAC TAAGACTTTC
61301 ATAAAAGCAC TGTATCTTAA CTTTCTGGC CTAAATCAAA AAAAGGAAAA
61351 CATTGATAAG TGTCTAGAA ACTTGGATTC TTTTATAGAT TTGTCTTGG
61401 GGCTCTGATG TTTGGGATTG ACGTCTCTGT CTGACCATTT TATATGCATT
61451 TTATCTTAAT AGTATGTGCT TTCATGAAGA TTCTGATACA AGTGGGCAAT
61501 CCTTAAATTA TCTTTGAAAA ATTGGTTAAT TTTGGTTAAA AAAGGGAAAG
61551 TGGCTGGGTG CAGTGGCTCA CGCCTGTAAT CCCCAGCACT TTGGGGAGGC
61601 GGGAGGGGTG GATCACAAGG TCAGGAGTTG AAGCCCATTC TGGCCAACAT
61651 GGTGAAACCC TGTCTCTACT GAAAATAATT GGGGCATGGT GGCACATGCC
61701 TGTAATCCCA GCTACTTGGG AAGCTGAGGC AGGAGAATTG CTTGAACCGG
61751 GGACCCAGGA GGCGGAGGTT GCAGTGAGCT GAGATCGCGC CACTGCACTC
61801 CAGCCTGGGC TACAGAGCGA GACTCTGTCT CAAAAAATAA ATAAATAAAT
61851 AAATGAAAAA GAGAAAAATAT TGAGAGGATT TGGTCATCAT TTTACTGCTC
61901 TCTTCATGTG ATGGAAATCA ATTTTCTCTC TCAAATGGGA TCAGTATCAT
61951 TTCCTAGTCA TACATCCATC CAGTTTTTGT TACTTTTTTG TTGGCATACA
62001 TTAATCAAAA TAGCTCTGCT TCATTGAGGC ATGCAGTCTG CAGACTCTCG
62051 GTGGAAAGGC TGTCATACTA TTAGTGACCA TAGTAACCTT TTATACCAAA
62101 GGATGGTTGC TGGATAATTT TAATATCTTT ACCAATAAAG TACTTTTTTG
62151 AAATACAAAA TCAGGCTGCT TGCTTTGCTC TATTCCTGTC AACAAAAAGG
62201 ATTTAGCTAT AGATTTAGCT TCTCCTTTTA TTTTCCCTTT TATTTCATAG
62251 GAGTCTTCTG TTTATTCCTT TCAGGCGCCT CCTTGGCATT ATAACAAAAA
62301 AAGATATCCT CCGGCATATG CCCCAGACGG CAAACCAAGA CCCCCTTCA
62351 ATAATGTTCA ACTGAATCTC ACAGATGAGG AGAGAGAAGA AACGGAAGAG
62401 GAAGTTTATT TGTTGAATAG CACAACTCTT TAACCTGAGG GAGTCATCTA
62451 CTTTTTTTTT CTCCTTTTACA AAAAAAGAAA GGAAATATAA AAGCCGGGTT
62501 TTTGCAACAT GGTTTGCAAA TAATGCTGGT GGAATGGAGG AGTTGTTTGG
62551 GGAGGGAAAG GAGAGAGAAG GAAAGGAGTG AGGTATTTCC CGTCTAACAG
62601 AAAGCAGCGT ATCAACTCCT ATTGTTCTGC ACTGGATGCA TTCAGCTGAG
62651 GATGTGCCCTG ATAGTGCAGG CTTGCGCCTC AACAGAGATG ACAGCAGAGT
62701 CCTCGAGCAC CTGGCCTGTT GCTCCAACAT TGCAAGACA CATTATCAGT
62751 CCCTATTTCT AGAGGGATTA CTTTGAATTG AGCCATCTAT AAAACTGCAA
62801 GGTCTTGCCC TTTTTTTTAA TCAAACTGT TCTGTTAAT TCATGAATTG
62851 TATAGTTAAG CATTACCTTT CTACATTCCA GAAGAGCCTT TATTCTCTC
62901 TCTCTCTCTC TCTCTCTCTC TCTCTCTACT GAGCTGTAAC AAAGCCTCTT
62951 TAAATCGGTG TATCCTTTTG AAGCAGTCCT TTCTCATATT GAGATGTACT

FIGURE 3, page 20 of 27

```

63001 GTGATTTTAC TGAGGTTTCA TCACAAGAAG GGAGTGTTC TTGTGCCATT
63051 AACCATGTAG TTTGTACCAT CACTAAATGC TTGGAACAGT ACACATGCAC
63101 CACAACAAAG GCTCATCAAA CAGGTAAAGT CTCGAAGGAA GCGAGAACGA
63151 AATCTCTCAT TGTGTGCCGT GTGGCTCAA ACCGAAAACA ATGAAGCTTG
63201 GTTTTAAAGG ATAAAGTTT CTTTTTGT TCCCTCTCAG ACTTTATGGA
63251 TAATGTGACC GGGTCTTATG CAAATTTTCT ATTTCTAAAA CTACTACTAT
63301 GATATACAAG TGCTGTTGAG CATAATTAAA TAAATGCTG CTGCTTTGAC
63351 AGTAAAGAGA AGGAAGTATT CTGATTAGCT GTATCTGGTA TTAATTGCAT
63401 GTTAAACAC TGGAATTTT AAAATTGAAA TTAGATCAGT CATTCCTTTC
63451 TTTTCTCAAG ATATCTCATG GCTGACACTG AAGAAGAAAT GTAATTCATA
63501 ACTTGCACTA AATGTATATT TTTTCTTCTA AAAATTACC ATTCTTATTT
63551 ATATTTTAT GGAATTAAT TTATAAATA CAGATCAGT AATATTGCAC
63601 TTAAGTAATT TTACCTTTT AATGTGATTT TTATAGAATA ATTCAGACTT
63651 ACAAATACAG AGATATGAAC AAAGTTTACA GTGGGAACAA AGGTTTAAAA
63701 AAAGGTTGTG GTTCTCTCTC TGTGATCCAG TGTGCACATA AACCTTCTC
63751 TGATCTTTCA CTGCCATCCT CTGGATTATG TCTTCTGACC TGTCATTTT
63801 GACCCATTAA CTGGAAGTT GAAAACTAC ATTAACGGGA AAGTTGAAAA
63851 ACTACATTAC TTTGGAGAAT AAAACCGAAA GTTCGTGTAT ACCTTCTTAA
63901 AAAAAAATC AAACCAAAAA TGTGAAAAA ATAGAATTGC AAAGATAGCA
63951 GTTAAATTT TAATCTGAAA ATAACCTTTG AATCTCGGGC TAGGTTATGT
64001 CCATATTTGA AGTGGTCAGT GATGTTTGA ACATTTTGT CAGGATGAGT
64051 TAAATGCAC TGGATTATAT TTGGGATTTT TGTTTTGGGA ATTGTCTGTT
64101 TTAATCACAG CCTTAATTCA CAATTGGCAA AGGCAGTTTA CTCAAAGGAC
64151 TGGGCTAAAT ATTCTGTAAT TATGCATTTT TGATAGGAAA ATGAAATTTT
64201 TGCAAAACAGA CATTTTCTTT TTTTGGCT GGAGTGCAGT GGGGCATGGT
64251 CTTGGCTCAC TGCAGCGTTG ACCACCTGGG CTCAAGTGAT ACTCCGCCT
64301 CAGCCACCA AGTAGCTGGC ACTACGGGCA CACGCCACCA TGCCAGCTA
64351 ATTTTTTGT ATTTTGTAGT GAGATGGGGT TTTGCCATGC TGCCAGGCT
64401 GGTCTCAACT CCTCAGCTCA AGCAATCTGC CTGCGTGAGC CTCCCAAAGT
64451 GGTGGAATTA CAGGCGTGGG CCACTGCGCC TGGCCAGAC AGACATTTTC
64501 TGAACACAA CTGGCAATGA GCTGTTTTTA CATTTTGAAA GTGATTCTTC
64551 ACTTCCTAGT TCTTAATTAT AGTATACCTA TTAAGATCTG TAAGATCCTG
64601 AAGACATAAG ATCATGAAGC CATATAAGAA TGAGGATTGA AAGTTGAGCA
64651 AAATTTTCGG GATTTTGGGA AACATCTTCA GCTGTGCTAT CTGCCTAAAA
64701 TTATTCCTTA TTACTTCTCT CCTTGACAG ACTTCAAGTT TTCTTCATAG
64751 CCCTTTCAAA GTTTTTTGAG CCATCCAGAG TAAATCATT TCTAAATGAT
64801 AGTTCTGTAT ATCTCCAAT CGTCTTAAGT GTATTGCCT GTGTGCAACG
64851 TATTGCTAGA CTATGAACTC CTCAGCATGG CTGCTGGATA ACTTAATTGT
64901 CCTGAGTTAA TAGCCTTCAA AGGACAAATC GGTTCCTTG CAGATAGCTT
64951 CGTAAACTT CACATGGAGT TTATTTTATC ATATTTCCCT TTTTATTTTC
65001 TGCTCCTCCT TTAATGCCC ATCTTGCTTC AGAGACTGAC ATTTCAGGGT
65051 GGATATTAAT TAAAGCATT AATTGTGTTT TTGGTATATT TCTATCCCTA
65101 GTATTTCTAT CTTACTGCTA AAATACAGGA AAAGTGCCGT ATTTTAAATG
65151 CATTTAGTGG TTTCTTTGG TGTATCTGT TCCATTTTTC TTTTTCATAC
65201 ATTGAAGTGT GTCTCCTTT CAACCAAAAT AATGAAATAG TGGAGACCAT
65251 GAAATGTTG TGCCTGGCTA ATTGGCAAAT TAATTTACCA ATATAATAAG
65301 TGTAGCGCCT GTTTTGAATA CCCTTTTGA GAAGGTATGA TGAGAATGGG
65351 CAAGGTGT (SEQ ID NO:3)

```

FEATURES:

```

Start: 2159
Exon: 2159-2237
Intron: 2238-22041
Exon: 22042-22199
Intron: 22200-30359
Exon: 30360-30459
Intron: 30460-31475
Exon: 31476-31663
Intron: 31664-32964
Exon: 32965-33087
Intron: 33088-34548
Exon: 34549-34755

```

Intron: 34756-37975
 Exon: 37976-38056
 Intron: 38057-39552
 Exon: 39553-40098
 Intron: 40099-46366
 Exon: 46367-46553
 Intron: 46554-49237
 Exon: 49238-49636
 Intron: 49637-55445
 Exon: 55446-55662
 Intron: 55663-62274
 Exon: 62275-62362
 Stop: 62363

CHROMOSOME MAP POSITION:
 Chromosome 4

ALLELIC VARIANTS (SNPs):

DNA			
Position	Major	Minor	Domain
1275	T	C	Beyond ORF (5')
1456	T	C	Beyond ORF (5')
5893	G	A	Intron
6226	A	G	Intron
8866	T	G	Intron
10397	C	T	Intron
10621	T	-	Intron
19651	A	G T	Intron
19891	T	-	Intron
20272	C	A	Intron
20412	T	A	Intron
23340	A	G	Intron
29948	T	A	Intron
33579	A	C	Intron
40762	G	A	Intron
40936	T	C	Intron
45998	A	G	Intron
47771	T	C	Intron
48117	C	T	Intron
54563	T	G	Intron
58735	C	T	Intron
59643	C	A	Intron
61638	G	T	Intron
63291	G	C	Beyond ORF (3')
63463	A	G	Beyond ORF (3')
63636	G	A	Beyond ORF (3')
63998	T	C	Beyond ORF (3')

Context:

DNA

Position

1275 GCATTTTCAGGAGGAGAATCTCCCAGTCTAGAGGAATCCTCTCAGAGGTAGCTATAAAATA
 TTGAACTCTGATCTTCAATAAGCATTGTGCGGTTTTTGTTTTGTTTTTAATGACAGTTT
 TAAACAAGAAAGTTGCTTTTATTTCTGAACCTTCATAAAATTTCTATTAAAGAGACAATTT
 CTGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTGTCT
 AAAAGTTAAAAACACAAAACCCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAG
 [T, C]
 CCTCTGCTTAGTAAACCTCTTTTTTGCGTAGTTTAGACACATACAATAGTAAAGTTACTT
 AGTACGTTGATAGTTTTCTTTCTCCTCAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
 AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC

TCTTCAGTTATTTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTT
TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTT

1456 TGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTTTGCTA
AAAGTTAAAAACACAAAACCTTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAGT
CCTCTGCTTAGTAAACCTCTTTTTTGCCTAGTTTAGACACATACAATAGTAAAGTTACTT
AGTACGTTGATAGTTTTCTTCTCCTCAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC
[T, C]
CTTCAGTTATTTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTTT
TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTTT
TTTTATGGCTATGACTCACATGACATTACTGTATAAACTAGTACATTCTCTCGTAAAC
CACACAACTTACTAGAGTGCTGCTCTCATTTTTCTACATTAGAAATGAAAAGGGCATT
GTCTGCATTCAAATTTCTTTTTACATCTCTGTATTACTTTTTCCCTTTTATATTTATC

5893 TCTAGTTGACAAGACTGAGGTAAGGAATTGTTAAGGAAAAGTCAGAATTCATCCAGATA
TTTGGCTCATACTTTAATCATGAGGCTAACTGCTTCTCTCTACAGTATCTTCATAGTA
ACTTGTGTTTTAAGTCTGGTAGAAGCATAAGAAGTTTAAACACAGACAGAATCCTGTGGA
AGTTAGTAAATTTCTAGTGAACGATAGAAATGATAGAAATCTCTTCTTCCCCAAAGTCC
CAAGAACAGATTAGTCTGCTTTTGACAAGTGTATCAAAGTAGACTGTTCTCACATACAC
[G, A]
GGGGACTCAATAGGGCATTCTCGGTGGATATAATAAAATGAGTAAATGCGATAACAGGAG
GAAATGCCTAGTGTGTTGCTCTTGGATTAGTTTTGATACAACAAAGGCAGCTTTGTTGTG
AGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGAAGTTGGAAGAGTGGCAGATCCTAGAGGA
CTAATGATGGGCTTAAACCACAAAAGTGTGCGTTTGCCATTGAA

6226 ATAAATGAGTAAATGCGATAACAGGAGGAAATGCCTAGTGTGTGCTCTTGGATTAGTT
TTGATACAACAAAGGCAGCTTTGTTGTGAGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGA
AGTTGGAAGAGTGGCAGATCCTAGAGGACTAATGATGGGCTTAAACCACAAAAGTGTGCG
CTTTGCCATTGAAATAAAAGTTTGGGGTCTTATTTTTTCAATTTTCTCCCTGAAATTATT
TCTTGACATTCATTAGCTCAGCAGTGTATCTAAATAAAGCTTTTTTGGGTTTCTATTATA
[A, G]
TAGAGGTTTGTTCCTTTTTCTTCCCTTTGAAAAGTATCATTTTTTGCACATTATTTGAAA
ATCCAGGTGTTATATGATATTTCTTATTGCCAGAGGGACATTCTGCAGGCTCTTTGTA
TGATTTTAGGATTCAGATACTTATTATATTTTTATTGGCCCTAATATTTTATCCAAC
TAAAGATTCTTTATTTGGTGTGAGAACTCCTTGTCTTACAAACAGTAGTATAAAACAA

8866 ACATGTAAACCAACAATGAAATTATTTTAGTGACTTGAGAATCAAAGTGCTAGAGTTTGA
ATCCCTGTTCTACTACTTGCTAGCGGTGTGACCTTGGGCTGTTTAACTCTTGACACCTT
GTTTTCCAAATTTATAAAGTGGAGATAATAATATCTGTACATTGTGTTGTTGTGAGGAT
TATATGAACATAATATGTAATGTCCTGAGAACAATGTCTGGTACACATTAAAGTTAATTA
AAATTAGCTGTTCTTACTGTTATTATTAGACATGAGCTAGATAACAGTGGCCTCTACATG
[T, G]
GAAAGATTATTTTAAATCTGATGTAGTTCAGTTTATCTATTTTTTTTTATTTTTGTCCCTT
TTGCATTGATGTATATCTAAAAACCTGCCTAACTCAGGATCACAAAATTTACTCCTG
TATTTTATAATTTTAGCTCTTTAGATCTAGGATCCATTTTTTAGCTAATTTTTATATATGG
TGTGAGGTAGGGGTACGGTTTCATTCTTTTGCACGTGAATAGCCAGTTGTCCAGCATCA
TTTATTCAAAGACTATTCTTTCTCACTAGAAAAATATTTCTTTAAAGAATAATGAAT

10397 CCAGGCTCCCTTGAACCTCCTGGGCTCAGATGATATAGCCTCCTGCCACAGCGTCTGATT
AGCTGGGACTACAGGTGTGACCACTACACGTGGCTTTCTGATGAAATTTTAAATACCC
AAATATTTGAGCAGAAATAATAGCTTGTGTTTATTGTTTTTCTACTATCTGTCAAGTATA
GTATTAAATGTTTACATAATTTGTCTCCAGTCCACATACTCTAGTAGAAGTGGG
TAACAAAACCAAGGTAATAAGGTTAATAAGTAACCTGCGCTGGATCACAGAATAA
[C, T]
GGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCAAAGTGCAGTAAAGTCAT
GGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGATCTTTTTTCCAGTCTTTTTGTTA
CTGTTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAGTAGTCACTGGCATCCGG
TAGTCAGCCCTCCAAAAAGTTTTTGATTTTTTTTTTTTTTTTTTGTCTTAACTTGAAG
CTACTAACTTTAGGTCATACTTTCTTATCATCAAGAGCTGGATATTTAGGTAGCAGAA

10621 CTCTAGTAGAAGTGGGTAACAAAACCAAGGTAATAAGTAACCTTGC

FIGURE 3, page 23 of 27

TGGATCACAGAACTAACGGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCA
 AAGTGCAGTAAAGTCATGGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGATCTTTTT
 TCCAGTCTTTTTTGTACTGTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAG
 TAGTCACTGGCATCCGGTAGTCAGCCCTCCAAAAAGTTTTTGATTTTTTTTTTTTTTTT
 [T, -]
 GTCTTAAACTTGGAAGCTACTAACTTTCAGGTCATACTTTCTTATCATCCAAGAGCTGGA
 TATTTAGGTAGCAGAACTATGGAATTATCCTAAGTCCTCTTGAAGCTTCAGCTGTAA
 ATTAATTGGTTCTGATTAACACTGTGCTCAAGATTACATTTCTAGGAGCCACAGTTGA
 TTGGTCTAACTTGGATCTATGTGTTTTCTTTAGCTGGGGAGGAGAAGGTATCTTGATTGA
 TACCTTCACCAGGACTGCATGCAGTGAGGGACAGAAGTTTCCTTAAATAATTGGGTTCT

19651 TTTATTTTCTGCTACTATGGCAGAATTGAGTTGTTGCAACTGTGTGGCATCCAAAGCCTA
 AAATATTTACTCTCCTGGCTCTTTGCCAACCCGTTTTAGATTATGAGCACTTTGGCATT
 TTATGTTTTTGTCTTTCTATAGCACACAGTAAGATGTTCTGCCACATTGTGCATAA
 TTTATGGGTTTATCAAGGATTTATGCAAGTGTAGCTGCAAGAAAAAACCTAGAAGTGA
 ACTTGCTAGGTTGAAGAGCA
 [A, G, T]
 CTGTGTATGTTAAATTTTGTAGCTTTTCGCCTTCCAAAGGGATTATTCCATTTCTACT
 TAACTACTAATTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATT
 ACACACTTCATCTTTACTAATCTGATAGAGGAAATGATATTGTGGATTGATTGTCATT
 TCTTTTATGTGTAGCTTGAGCTTATTTTCATATTTAAAGCCAATTGTATTCTTTTT
 CTTGAGCTATCTTTTAATGT

19891 TTTATGCAAGTGTAGCTGCAAGAAAAAACCTAGAAGTGAAGTGTAGGTTGAAGAGCA
 TCTGTGTATGTTAAATTTTGTAGCTTTTCGCCTTCCAAAGGGATTATTCCATTTCTAC
 TTAAGTACTAATTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATT
 CACACACTTCATCTTTACTAATCTGATAGAGGAAATGATATTGTGGATTGATTGTCAT
 TTTCTTTTATGTGTAGCTTGAGCTTATTTTCATATTTAAAGCCAATTGTATTCTTTTT
 [T, -]
 CTTGAGCTATCTTTTAATGTCTTCTCTGATACATTTCTGAAGTCTGTGATACTCATATAA
 GATATATGGTGAACATGTGTCAAAGATTTATTTGACTCTAATGAGGGAACCCGCTGATG
 ACAAGGCTGATTGAGAAGAGGATGTGTGAGATGAAGTGTATATCATCAGTGAAGAAAGC
 AAATCTTACAGGGCAAAAACAAACCACTAAGGGTTATTGTTTCTACTGGACAG
 AATTCATTTGCATTTTACCAGATAAAAATTACTATTTTCAATTTATCTTTTACAAATCAT

20272 CAAAGATTTATTTGACTCTAATGAGGGAACCCGCTGATGACAAGGCTGATTGAGAAGAG
 GATGTGTGAGATGAAGTGTATATCATCAGTGAAGAAAGCAAATCTTACAGGGCAAAAA
 CAAAACCACAACCTAAGGGTTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCA
 GATAAAAATTACTATTTTCAATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCT
 ATTCCCTAATCAGTAGTAAATAGTCTTCAAATTTCTCCGAGCGTCAGGTGACTATTATG
 [C, A]
 AGGCTAATTGTTGACACTCGGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTT
 ACTTCCAAGTTTTGGATAATTTTTCTTAACACATTTTCTCTAATTGCAATGATTTCAAG
 TGATATTATTTCTTTTTTAAATTTTTTACTATTTATTGATCACTTTGGGTGTTTCT
 CGGAGAGGGGGATTGTCAGGGTCATAGGACAATAGTGGAGGGAAGGTGAGCAGATAAAC
 ATGTGAACAAAGGTCTCTGTTTTCTAGGCAGAGGACCCTGCGGCCTTCCACAGTGTTT

20412 TTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCAGATAAAAAATTACTATTTTCA
 ATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCTATTCCCTAATCAGTAGTAA
 TAGTCTTCAAATTTCTCCGAGCGTCAGGTGACTATTATGCAGGCTAATTGTTGACACTC
 GGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTTACTTCCAAGTTTTGGATAA
 TTTTCTTAAACATTTTCTCTAATTGCAATGATTTCAAGTATATTATTTCTTTTTTT
 [T, A]
 AAATTTTTTTTACTATTTATTGATCACTCTTGGGTGTTTCTCGGAGAGGGGGATTGTCAG
 GGTCAAGGACAATAGTGGAGGGAAGGTGAGCAGATAAACATGTGAACAAAGGTCTCTGG
 TTTTCTAGGCAGAGGACCCTGCGGCCTTCCACAGTGTTTGTGTCCCTGGGTACTTGAGA
 TTAGGGAGTGGTGATGACTCTTAATGAGCATGCTGCCTTCAAGCATCTGTTTAAACAAAGC
 ACATCTTGACCGCCCTTAATCCCTTTAACCCTGAGTTGACATAGCACATGTTTCAGAGA

23340 TTTTTTTTTTTGGAGGTGCGGGGACTGTGCCCCATTCTGTTGCCCAAAGTGGAGTGCAGTG
 GTGCAATCTTGGCTCACTGCAACCTCTGCCTCCAGGTTCAAGCGATTCTTGTACTCAGC
 CTCTGAGTAGCTGGAATTATAGGTGTGTGCCATCATGCCAAGCTAATTTTTGTATTTTT
 AGTAGAGATGAAGTTTCGCCATGTTGGCGAGGCTAGTCTCAGACTCCTGGCCTCAAGTGA

FIGURE 3, page 24 of 27

TTGGCTGACCTCAGCCTCCCAAAGTAGAAAATCTTCTTGAAAAATAAAATTCCAAATCTC
[A, G]
AAAGGCCCTATATAATTTTGGTGTGGAAATTTACTTGTCAATGAAATGACTATTTACA
CAAATTATAAGCTTCCATATTTAAATATATATGTGTGAACCTGAAATTCAAATTTATTA
TATTGTTTATGAAAGGTACAGCCTCTGAGATTCATCAGATGGTATTTACCTTTAGGGCAT
ATCTAAAAATAAAATACAGTACATGAAATCCAGTGCTTTAATCCAGTGATTCTTAAACTT
TTTGCTCTCAGATCCCCTTTAAACTCTTAAAAGATATTGAAGAGCTCCAAGGAGGCTTTG

29948 GACTCTACCAATGGGATCGGAGCTCTCCAAACCTGCATATTAAGGCCTATAAGTTTTG
GGGGTCCCTTTGTCCACATGATTATTCTGTAATACATTGTATTATGGACATGGTATTA
TTATACACAGATCCTGTCTTTTAAAGAACATTATAATCCACTTAACTGCTAGGACGAG
AATGACCGATAATTCAAACCATATTGTCTTACAGAAGACATATATAAAGATGGTTATGT
GTACCAATTGAGGTTCAAATTTGATTCAATTTAAAACAATCTAGGCCAGATTTTATATAG
[T, A]
TTGTGGACCTTTGCACTCAAATCTCAAGGTTCTTATTAAAATGCAGATCTTGGCTGGGC
ACGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGCCCAAGGCAGGTAGATCATTGGA
GCTCAGAATTCAAGACCAGTCTGGCCAACATAGCGAGGCCAGTCTCATTGAAAGAAAA
AAAATTTTTTAAATAAAAAATAAAAGCAGATCTTGGGTAAAGACATGTAGTCTGGTTTACA
GGTATTAACAACGTCTGTAATGTAGTGATTTTGCTCCAGACTTACCTTTTCCATTATTT

33579 TTACTGTGAAGGCTGATTTTTTTTTTCTCTCACCCTAATTTAACACATGACTAGGCAA
TTTTTCAGACTATTTAGTTAAACATCAAGAGCCTGGAAGAAGTATCTTGTGACCTAATGTT
CTTTGACGGGTTAGTTGTTACTTTGCTGTTATGACCTGAATTTTTTTTTTTGAGACTG
AGTCTTGTGCTGTGCGCCAGACTGGAGTGCAGTGGCGCAATCTCAGTCACTGCAACCTC
TGCGTCCCAGGCTCAAGCAATTTCTGTGTCTCAGCCTCCTGAGGAGTTGCGATTGCAGGC
[A, C]
CCTGTCAACATGCCCTGCTAATTTTTGCTTTTTTGTGTTGTTTTTTTTTTAGTAGAG
ATGGGGTTTACCCTGTTGGCCAGGCTGGTCTCAAACCTCCTAAGTATCACCCTG
CCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACAGTGGCTATGACCT
GATTTTGATTCACTCTTTTATAATTACCTTTTGATTAGATAAGTTAATTATCTTGA
ATTTGGCCATTTATGCTTTGAGAAAGTAGTTAATCACAGTGGGTCAACAGTACAACTT

40762 ATCCATCACCTCAAGCATTTATCCCTTGTGTTACAAACAATCCAATTACACTCTTAATTA
TTTTTAAGTGACATTAATTAATTATGAATATAGTTCAAAGACTTCTTCATTGACTAG
CACCTAGGCTAAAAAATTCAGACACCTGGGCTCCTGGGATCAATCACGCATCTGTGTC
TCTTGTGCTCACTCCC
[G, A]
CTGTCTCTCTCTCTTTCTCTCGCTTCCTTTTTCTCTCTCTCTGTTTCTAGGGTGG
TGGCCTCAGGGAATTGGATTCTTATATTATAGCTCAGGATTCCCAAGAGGGCTGTTTT
AATGTAGCCAAAGAAGTCTTGACGCGTGAATTTGTTTATTCTATTGAGGTAGTCAC
AGAGGCCCGACCACAT

40936 TACAATTAATTTATGAATATAGTTCAAAGACTTCTTCATTGACTAGCACCTAGGCT
AAAAAATTCAGACACCTGGGCTCCTGGGATCAATCAGCATACTGTGTCTCTGTGCTC
ACTCCCGCTGTCTCTCTCTCTCTCTCTCGCTTCCTTTTTCTCTCTCTGTGTTTTCT
AGGGTGGTGGCCTCAGGGAATTGGATTCTTATATTATAGCTCAGGATTCCCAAGAGGGC
TGTTTTTAATGTAGCCAAAGAAGTCTTGACGCGTGAATTTGTTTATTCTATTGAGG
[T, C]
AGTCACAGAGGCCCGACCACATTCAGAGGAGGGACATACACTTGCTGGGACAAGTGAAG
AGAATTCATGATCATGTTTTTAAACCACTTTTATAGTTTCTATTGCTGCTGTAATAAA
TTACCAACAATTAATGGCTTAAAGCCACACAAATTTAATATCTTACAGTTCTGCAATC
AAAAGTCTGAAACGGATCTCACTGTGCTAAAATTAAGGTGTTTCGTAGGGCATTCTGGAGG
CTGTAGGAGAGAGTCTGTTTTTTGCCTTTTCTGGCTATTAAAAGCTGCCAGCATTCTT

45998 TGTATATCAGTCAAAATATTGGGCAACTCTGATAAGTTTGTCCACTTAACATTGTACCAC
TTAAGATGAATAGCATCTACCATTTCCGTCATTTGTAAATATATAGGAGGACATAATCAC
ATAATCTTGAAGTAAAGACAGTGCTTAAACTGAATCAGTTAAGTTTTATGAAAAATAC
TTATATTGTACTTTTAAATAATATATATTTTTTAAATTTCAATAGCTTTTGGGTACAAAGT
GGTTTTGGTTACGTGGATGAATCTATAATGGTGAAGTCTAAGATTTTACTGCAACTGTC
[A, G]
CCCAAGTAGTATATATTGTATCCAGCATATTGTCTTTTTTTTTTCTTTTTTTTTTTTCA
TTTACCATGGACTAATGAAATTTTGTAGGGACTGACATTAGGGCACCTTGAGCTAC
CTTGAGCTAAAGGAAATAACCTTGAATTTTTTCTGTTTGGCCTAGAGAATGTGGTTTG

FIGURE 3, page 25 of 27

TTTTGTAAGTGAATTCATGGGATTGTTAAGGTACAAGATTTTGCTTTAGTTTTATTTGTA
CTAGGATTTTGCTATATTAATACAATGTGAAAAGAATCAAAGTGTTAGAAATAAATGCA

47771 GAAGAGTAGAACATGAGGCTTTATTTAAAAGATTAGCAGAATTTAAGGAAAAGGTGACTT
TGTTGAAGATTATAATGTGAAGACAAAGGAACGAGGATGGGAATAAATTTTGATTTCATG
AGGCTTTGAAGAAATTGACTCTAGAGAGTATATTTGGGTACTTTGGGAAATGAAGTTG
GATTAGTGAGAAGGAACAGATTATGAAAAGACAAGAAACCTGATTAATGTCAGGATGATT
TTATATTTGAAG
[T, C]
TGGTCAGATTTATGGCAGTCCTGGCTTTGCCATTTTGTAGTTTGATGACTTTGAGAAAAGTT
CCTTCTTGAAGTTTAAATTTTCTGTATATAAAAAGTAATAACACCTGGTGATCTGCTAGG
TTTGTGTTTGAAGATTATATGAGATAAAAATGCATGCAAACTGTTATAATAGTGCCTGGTA
AAATAAGTGCCTAGTTTAAAAACAAGTCTTTGTAACTGCTTAGGACATGCCTGGTATA
GGGTAGGTATGT

48117 GACTTTGAGAAAGTTCCTTCTTGAAGTTTAAATTTTCTGTATATAAAAAGTAATAACACC
TGGTGATCTGCTAGGTTTGTGTTGAGGATTATATGAGATAAAATGCATGCAAACTGTAA
TAATAGTGCCTGGTAAAATAAGTGCCTAGTTTAAAAACAAGTCTTTGTAACTGCTTAG
GACATGCCTGGTATAGGGTAGGTATGTAATACATAGTAGGTAGGATCTGTCTCCTTGCTA
TTTTTAGGTAAAAAACAAAAGGAAGAGCTTCAGCTTAATACAGTATGAAGTACGAGGCC
[C, T]
TGGTAGGTTTTTTGAGCAAAAGAGCAACACAGTAAAAGTAGTACTTAGGAAAGATTAACAA
GGGAACATGGCTTATACAGTGGTAATGGGGCCTGGAGTCAAGGAGGTAAAGATAAAATGGT
ATTATAATTAAGGAATAGCCAGGCACGATGGCACATGCATGTAATGCCAGCTACTGGAGA
GGCTGAGGTGGGAGGATCATGGGAGTCCAGGAGTTTGAGACCAGCCTGGGCAACTGAGTG
AGACCCCAATCCTAAAAAATACAAAAGTAAAAAAGGAATAAAGTCATGAGGGCTTGACT

54563 GCTTTGTCACCCAGGCTGGAGTGTGGTGGTATGATCATGGCTGACTGCAGCCCTGACCTT
CCGGGCTCAAGTGATCTTTCCACCTCAGCCTCCCAATTACTTGGGACCACCAGCATGCTT
GGCCGATTTTTTTTTTTTTTTTTTTTTTTTGTAGAAGCAAGTTTCCCTATGTTGCCAAGGC
TGGTCTTGAACTTAGGGCTCATGTGATACTCCTGCCTCGGCCCTCCCAAAGTGTAGGAT
TACAAGCCTGAGCCACCATGGCCGGCCAAAATATTTTCACTATAACAAATATCATATCTG
[T, G]
ATATACTCAGTTTTAATACTAACTCAAAGTAGAAACATAAAGCTGAATGACTATTTTTATT
TTCAGATTCTCTCCATTGAGTTTCCTTCTCCGTCTGTGTGATCTCTGAACTTTCTCCA
TCTTTGCCACTTCTTGTCTAGCATTTTTTTTTTATCAGCAGTTTCATTTCAGATTTTTTTT
TTAGTTCTTTCAACGGTGGAGTGGAAAGTAGGCAGCAGGACAGAAGAACTTGAAGCAGAGC
ACACTGGAGAGGAGAAATTAACAAAGCCTTTATGAATAAAACAACCCCCCAATATCAGTC

58735 TGGGTTATGCCCTGTTAACTCTTACATCATTAGTTTTTAGCCCCAAAAGGAAACAGCAAAT
AATGTTTTATATGAGCCACATTTTGCCTTGATTTTCCCTTCCACTCTGTAAAATTAATAA
GCAGCACTCTGACTTTATTATGCTCAAATCGCTCTTCTCCATTAATGTGTGTTTCTCCAT
CTTTTAGGGTTTTTACTTTATAAATACAGAGATTACTGTGTAAAATTCTAAATTTGCCAC
TGGGTCGTTATACATTTGTAACCTTCCTCACAGTATATTTTGATTTGGCAGAGTTTAC
[C, T]
AATATAGATGATACTAACTGAAATTAATCATTCTGTATAATTGGATAGAAAAGCATGAGT
AAGAATTCAATTGGTATTATATTTAATTAATTGCCAAGATTTTCACATTTCTGACTACA
ACAATAAAATCAAATGAATTGATGGCTTAAAAAAAAGAAATCTCAAATGTTTAGTCAATG
AAGAACATCTATTGAATGAGTGAATGTTTATTATATAGTGCATTTTCTGAGCTTTTTT
GGAGGGGGAAGTTGCTCCCATGCTCTGAGAACTTTAAGGATCGATACATTATTTTAAAC

59643 GTTTATATTGCCACATTAATTTCCATTATAAAACCAGTAACCATAGTTTTGTTTAAATTA
GCAATCTAATTATTTTCATGTATCCTCATTATGAGAATTTATGTCCATCACTTTGCTTGA
TGTGATAACAGTGACATGCTAAATGAGAAACAATTGTTATTTAGAAAAAATGCACAAAG
TGAAAGTCCTTTTAAATCCCTAATCATAAATACATTTTATTAGCTTACTTTAAGAAGTGGC
AGTCACAGCTCCTGAACATTAGGGAGTGTTCCTTTGGTCAGCATTATTTATTTAGTGCA
[C, A]
ATTGCCCTTTAATTTTAAATTTGAAATTATAGTAAAATCCACGGGAGTTTTTAAGTCTCCTC
ACAGCCTTTTGCTACCTTTTCCCAAGGTAGATCCAGATGATAACTGCTGTGTGTGACA
TCATAGAAATTAGAAAAATATTTCTCTGAGGAAAGAACATTGTAAATGAACTCTACA
TATCAGAGGTCTATAGCTATGTATCAATATTAAGTTTCTTTTGTACTTTGCTTTGTAGTC
ATCTTCATTCCAACTTTTCAATTTATTATTTTACTTTAAAAAGAAAAATAACCCACCA

FIGURE 3, page 26 of 27

61638 AAAAAAAGGAAAACATTGATAAGTGTCTAGAACTTGGATTCTTTTATAGATTGTCTT
TGGGGCTCTGATGTTTGGGATTGACGTTCTGTGCTGACCATTTTATATGCATTTTATCTT
AATAGTATGTGCTTTTCATGAAGATTCTGATACAAGTGGGCAATCCTTAAATTATCTTTGA
AAAATTGGTTAATTTTGGTTAAAAAGGGAAAGTGGCTGGGTGCAGTGGCTCACGCCTGT
AATCCCCAGCACTTTGGGAGGCCGGGACGGGTGGATCACAAAGTCAGGAGTTGAAGCCCA
[G, T]
TCTGGCCAACATGGTGAAACCCTGTCTCTACTGAAAATAATTGGGGCATGGTGGGCACATG
CCTGTAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATTGCTTGAACCGGGGACCCAG
GAGGCGGAGGTTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGCTACAGAGC
GAGACTCTGTCTCAAAAAATAAATAAATAAATAAATAAAGAGAAAAATATTGAGAGGA
TTTGGTCATCATTTTACTGCTCTCTTCATGTGATGGAAATCAATTTTCCTTCTCAAATGG

63291 GAGATGTACTGTGATTTTACTGAGGTTTCATCACAGAAGGGAGTGTTCCTGTGCCATT
AACCATGTAGTTTGTACCATCACTAAATGCTTGAACAGTACACATGCACCACAACAAAG
GCTCATCAAACAGGTAAAGTCTCGAAGGAAGCGAGAACGAAATCTCTCATTGTGTGCCGT
GTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTAAAGGATAAAGTTTTCTTTTTTGT
TTCTCTCAGACTTTATGGATAATGTGACCGGGTCTTATGCAAATTTTCTATTCTAAAA
[G, C]
TACTACTATGATATACAAGTGTCTGTGAGCATAATTAAATAAAATGCTGCTGCTTTGACA
GTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCATGTTAAACACT
GGAATTTTTTAAATTTGAATTAGATCAGTCATTCTTTTCTTTTCTCAAGATATCTCATGG
CTGACACTGAAGAAGAAATGTAATTCATAACTTGCCTAAATGTATATTTTTTTCTTAA
AAATTTACCATTCTTATTTATATTTTTTATGGATTAAAATTTATAAAATACAGATCAGTTA

63463 TGTGCCGTGTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTAAAGGATAAAGTTTTCT
TTTTTGTTCCTCTCAGACTTTATGGATAATGTGACCGGTCTTATGCAAATTTTCTAT
TTCTAAACTACTACTATGATATACAAGTGTCTGTGAGCATAATTAAATAAAATGCTGCT
GCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATGCATGT
TAAACACTGGAATTTTTTAAATTTGAATTAGATCAGTCATTCTTTTCTTTCTCAAGAT
[A, G]
TCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTTGCCTAAATGTATATTTTT
TTTCTTAAAAATTTACCATTCTTATTTATATTTTTTATGGATTAAAAATTTATAAAATACAG
ATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTAAATGTGATTTTTATAGAATAATT
CAGACTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGGTTAAAAAAA
GGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTTCACTG

63636 TGCTGCTGCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAAT
TGCATGTTAAACACTGGAATTTTTTAAATTTGAATTAGATCAGTCATTCTTTTCTTTTC
TCAAGATATCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTTGCCTAAATGT
ATATTTTTTTTCTTAAAAATTTACCATTCTTATTTATATTTTTTATGGATTAAAAATTTATA
AAATACAGATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTAAATGTGATTTTTATA
[G, A]
AATAATTGCACTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGGTTT
AAAAAAGGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCT
TTCACTGCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAACTGGAA
AGTTGAAAACTACATTAACCTGGAAGTTGAAAACTACATTACTTTGGAGAATAAAACC
GAAAGTTTCGTGTATACCTTCTTAAAAAAAATCAAACCAAAATGTGAAACAATAGAA

63998 AAAAAAGGTTGTGGTTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTT
TCACTGCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAACTGGAAA
GTTGAAAAACTACATTAACCTGGAAGTTGAAAACTACATTACTTTGGAGAATAAAACCG
AAAGTTCGTGTATACCTTCTTAAAAAAAATCAAACCAAAATGTGAAACAATAGAA
TGCAAAGATAGCAGTTAAATTTTAACTGAAAATAAACCTTTGAATCTCGGGCTAGGTTA
[T, C]
GTCCATATTTGAAGTGGTCAGTGATGGTTTGAACATTTTTTGCAGGATGAGTTAAATGC
ACTGGATTATATTTGGGATTTTGTTTTTTGGAAATGTCTGTTTAAATCACAGCCTTAATT
CACAATTGGCAAAGGCAGTTTACTCAAAGGACTGGGCTAAATATTCTGTAATTATGCATT
TTTGATAGGAAAATGAAATTTTTGCAACACAGACATTTTCTTTTTTTTGGCTGGAGTGCA
GTGGGGCATGGTCTTGGCTCACTGCAGCGTTGACCACCTGGGCTCAAGTGATACTCCCGC

FIGURE 3, page 27 of 27